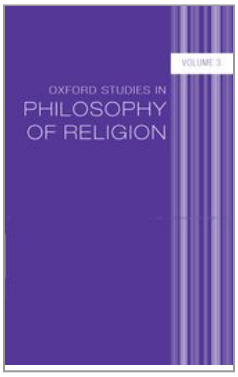


Oxford Studies in Philosophy of Religion Volume 3

Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

Title Pages

(p.i) Oxford Studies in Philosophy of Religion

(p.ii) Editorial Board

(p.iii) Oxford Studies in Philosophy of Religion Volume 3

ROBERT AUDI, UNIVERSITY OF NOTRE DAME
KEITH DEROSE, YALE UNIVERSITY
HUGH MCCANN, TEXAS A&M UNIVERSITY
BRADLEY MONTON, UNIVERSITY OF COLORADO
GRAHAM OPPY, MONASH UNIVERSITY
ELEONORE STUMP, ST LOUIS UNIVERSITY
LINDA ZAGZEBSKI, UNIVERSITY OF OKLAHOMA
TED A. WARFIELD, UNIVERSITY OF NOTRE DAME

OXFORD
UNIVERSITY PRESS

(p.iv) OXFORD UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide in Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York

© The several contributors, 2011

The moral rights of the authors have been asserted
Database right Oxford University Press (maker)

First published 2011

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any
means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the
appropriate
reprographics rights organization. Enquiries concerning
reproduction
outside the scope of the above should be sent to the Rights
Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

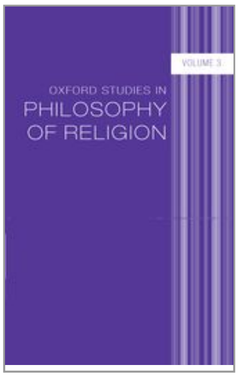
British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain
on acid-free paper by
MPG Books Group, Bodmin and King's Lynn

ISBN 978–0–19–960321–3 (Hbk.)
978–0–19–960322–0 (Pbk.)

1 3 5 7 9 10 8 6 4 2



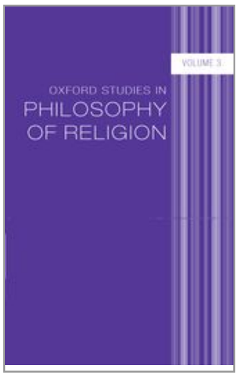
Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

(p.vii) Editor's Introduction

This is the third volume of the *Oxford Studies in Philosophy of Religion* series. As with the first two volumes, these essays follow the tradition of providing a non-sectarian and non-partisan snapshot of the subdiscipline of philosophy of religion. This subdiscipline has become an increasingly important one within philosophy over the last century, and especially over the past half century, having emerged as an identifiable subfield within this time frame along with other emerging subfields such as the philosophy of science and the philosophy of language. This volume continues the initial intention behind the series of attracting the best work from the premier philosophers of religion, as well as including top philosophers outside this area when their work and interests intersect with issues in the philosophy of religion. This inclusive approach to the series provides an opportunity to mitigate some of the costs of greater specialization in our disciplines, while at the same time inviting greater interest in the work being done in the philosophy of religion. We thus present the third volume in this series, a volume containing contributions by an impressive group of philosophers on topics of central importance to the philosophy of religion.

Jonathan L. Kvanvig **(p.viii)**



Oxford Studies in Philosophy of Religion Volume 3

Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

(p.ix) List of Contributors

Michael J. Almeida is Professor of Philosophy at the University of Texas, San Antonio.

Daniel Bonevac is Professor of Philosophy at the University of Texas, Austin.

E. J. Coffman is Assistant Professor of Philosophy at the University of Tennessee-Knoxville.

Phil Dowe is Associate Professor of Philosophy at the University of Queensland.

Peter J. Graham is Associate Professor of Philosophy at the University of California-Riverside.

Frances Howard-Snyder is Professor of Philosophy at Western Washington University.

Kevin Kimble is Assistant Professor of Philosophy at National Chung-Chen University, Taiwan.

Bradley Monton is Associate Professor of Philosophy at the University of Colorado, Boulder.

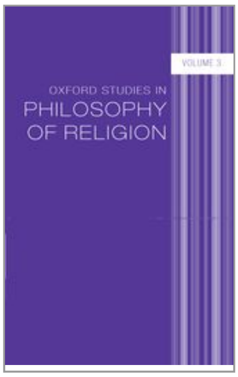
Timothy O'Connor is Professor of Philosophy and Department Chair at Indiana University.

Graham Oppy is Professor of Philosophy at Monash University, Australia.

Elliott Sober is Hans Reichenbach Professor and William F. Vilas Research Professor at the University of Wisconsin-Madison.

Patrick Todd is Visiting Graduate Fellow, Center for Philosophy of Religion at the University of Notre Dame.

(p.x)



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

Theistic Modal Realism? *

Michael J. Almeida

DOI:10.1093/acprof:oso/9780199603213.003.0001

Abstract and Keywords

Genuine modal realism and counterpart theory afford theists interesting ways to resolve some recalcitrant problems in philosophical theology. There are straightforward solutions to the problem of No Best World, the Modal Problem of Evil and the problem of the Less-than-Best World. But genuine modal realism affords a solution to even more difficult problems including how a morally perfect being might fail to prevent gratuitous evil and how a morally perfect being might perform a morally wrong action.

Keywords: modal realism, counterpart theory, modal problem of evil, no best world, gratuitous evil

I. Introduction

In good Leibnizian tradition Phil Quinn maintained that a strong improvability principle must govern the choices of perfect beings:

> If an omnipotent and superlatively good moral agent were to actualize a possible world he would actualize some…world of unsurpassable moral goodness.[1]

Since it is impossible for a perfect being to actualize more than one possible world, theists in the Leibnizian tradition are committed to the unlikely proposition that the actual world, with all of its evil, is as good as any other actualizable world.[2] Call that the *Less-than-Best Problem*.

One theistic response to the *Less-than-Best Problem* is to maintain that every possible world is a *real, concrete universe*.[3] Theistic modal realists might take the position that our world is simply one among an infinite plurality of concrete universes actualized in logical space. It makes no important moral difference that, from our point of view, our world is actual. If there is a best possible world, theistic modal realists can argue, there is no moral reason why *ours* should be that world. There is no moral reason why the individual inhabitants of our world should enjoy the best possible experiences rather **(p.2)** than the individual inhabitants of other real concrete universes. And if there are some bad possible worlds then there is no reason why ours should not be among them. In the “multiverse” that includes our universe there is the greatest total amount of value possible. In section 2, I briefly describe some features of modal realism. In sections 3–8, I argue that theistic modal realism has the resources to resolve a series of problems derived from the *Principle of Plenitude*, including the *Modal Problem of Evil* and the *Less-than-Best Problem*. I offer some concluding remarks in section 9.

II. Some Features of Genuine Modal Realism

Mark Heller offers this initial characterization of genuine modal realism:

> Modal realists…believe that the actual world is a concrete object of which you and I are literal parts, and he believes that other worlds are also concrete objects some of which literally include other people as parts. Merely possible worlds and merely possible people *really exist* despite their lack of actuality.[4]

Suppose there is an infinite plurality of possible worlds. Every possible world is a real, concrete universe and each world is a causally and spatio-temporally closed individual. None of the infinite plurality of possible worlds stands in any causal or spatio-temporal relation to any world other than itself. And no world stands in a causal or spatio-temporal relation to the parts of any worlds other than its own parts.[5]

We are parts of the actual world, or *world-mates*, because we stand in spatio-temporal relations to one another. And for any possible world *w* the individuals in *w* are parts of *w* because they stand in literal spatio-temporal relations to one another. All individuals are worldbound. No individual exists in more than one world. Among other things, this entails that our world does not overlap any other world with respect to people, quarks, leptons, water molecules, or any other part of the world. But for any individuals at any world there is some world containing duplicates of those individuals and many worlds containing counterparts of those individuals.

Suppose that at each concrete, spatio-temporally isolated universe there is an Anselmian perfect being that actualized that universe.[6] The Anselmian **(p.3)** perfect being necessarily exists, so each world includes a world-bound perfect being that has the property of necessarily existing.[7] For every valuable experience that an individual could have there is some individual that does have that experience in some world. And for every valuable thing that could exist there is some world at which that valuable thing does exist.

III. Plenitude Problems for Theistic Modal Realism

According to David Lewis's initial formulation of the *Principle of Plenitude*, absolutely every way that a world could be is a way that some world is and absolutely every way that a part of a world could be is a way that some part of some world is.[8] To express the plenitude of possible worlds, Lewis appeals to a *Principle of Recombination*:

> …according to [the principle of recombination] patching together of parts of different possible worlds yields another possible world. Roughly speaking, the principle is that anything can coexist with anything else, at least provided that they occupy distinct spatiotemporal positions. Likewise anything can fail to coexist with anything else. Thus if there could be a dragon and there could be a unicorn, but there couldn't be a dragon and unicorn side by side, that would be an unacceptable gap in logical space, a failure of plenitude.[9]

The *Principle of Plenitude* is supposed to ensure that there are no gaps in logical space. There is some real concrete universe for every way a world could be. Of course it is difficult to know exactly how many ways a world could be, but the plurality of worlds would presumably include some worlds that are on balance bad.[10] Otherwise there would again be an unacceptable gap.

Suppose then that *w* is an on balance a bad world. It is of course true at *w* that *w* is actual, and presumably the perfect being at *w* could have prevented *w* **(p.4)** from becoming actual. Anselmian perfect beings have all of the attributes of perfection including essential perfect goodness, essential omnipotence, essential omniscience, and necessary existence. Here is Theodore Guleserian:,

> Presumably, an omnipotent being has the power to prevent any possible world from becoming actual, since all one has to do to prevent a world from becoming actual is to bring about some state of affairs that is not included in that world.[11]

But if it is true at *w* that a perfect being either brings about *w* or allows *w* to be actual, then it must be true at *w* that a perfect being is morally permitted to actualize *w*. Anselmian perfect beings are essentially morally perfect and cannot perform any impermissible actions. But then an Anselmian perfect being is not permitted to actualize *w*. Here is Guleserian again:

> There is a possible world *w* such that necessarily [there is a perfect being in *w*] only if it is *not* morally permissible for [that being] to allow *w* to be actual.[12]

Theistic modal realism entails that each possible world is a real concrete universe that a perfect being has actualized.[13] But the *Principle of Plenitude* entails that at least some of those worlds are so bad that no perfect being could actualize them. Theistic modal realism is therefore inconsistent with the *Principle of Plenitude*.

IV. Plenitude Problems Reconsidered

Theistic modal realists might decide to abandon the *Principle of Plenitude*. Thomas Morris has urged, for instance, that Anselmian theists should conclude that there are no possible worlds that a perfect being could not actualize:

> …[An Anselmian] God is a delimiter of possibilities. If there is a being who exists necessarily and is necessarily omnipotent, omniscient, and good then many states of affairs which otherwise would represent genuine possibilities, and which by all non-theistic tests of logic and semantics do represent possibilities, are strictly impossible in the strongest sense.[14]

But the modal position Morris describes seems to beg the central question at issue. Even a modest position on the deliverances of modal intuition urges that **(p.5)** at least some possible worlds are on balance bad. Morris's position is that, for committed Anselmians, the otherwise credible deliverances of modal intuition are not reliable guides to what is genuinely possible. But one of the questions at issue is whether anyone ought to be a committed Anselmian in the first place. And certainly our considered intuitions should play a role in delimiting possibilities for those considering the possibility of an Anselmian God.

Theistic modal realists might urge instead that since God creates the multiverse—that is, every possible world in logical space—the sum total of value across the vast pattern of possible worlds must be *on balance positive* or, at least, *on balance neutral*. According to this view, there are many real concrete universes that are on balance bad and there are many real concrete universes that are on balance good, but the sum total of value across all possible worlds is positive; the sum total of all God has created is positive. And this position is consistent with the *Principle of Plenitude*. Theistic modal realists might then conclude that the existence of a perfect being is compatible with a multiverse that is on balance positive.

But a theistic modal realist must concede that the multiverse includes many spatio-temporally independent possible worlds each of which a perfect being actualizes. Some of those independent universes are on balance bad. Certainly theistic modal realists need some explanation of how a perfect being might actualize a universe that is on balance bad.

V. Plenitude Problems Resolved

Since the *Principle of Plenitude* entails that there are at least some worlds that are on balance bad, we are forced to conclude that the traditional Anselmian God exists in some worlds on balance bad. Let *w* be on balance a bad world. According to theistic modal realism, *w* is no different from the actual world in ontological kind. Both are concrete worlds containing various parts instantiating various properties. The pain and suffering endured in *w* is no less bad than the pain and suffering endured in our world. It is true at our world that the suffering endured is genuine, but the suffering endured at *w* is also genuine. The strong atheological claim is that no Anselmian perfect being actualizes any world like *w* that is on balance bad. Call the strong atheological claim SA.

SA. If an Anselmian God exists, then there are no on-balance bad worlds. To make the problem more concrete, suppose Smith is a moral agent in *w* and **(p.6)** Smith is suffering some affliction. Suppose it is true in *w* that Smith is a good and just person. He is undeserving of the suffering he has endured, and his suffering is preventable without moral cost. Smith's suffering seems to be a clear instance of gratuitous evil.[15] So, according to SA, it is false that the Anselmian God exists in *w*. No perfect being allows good and just moral agents to suffer undeserved and preventable afflictions.

A theistic modal realist should concede that a perfect being could have prevented Smith's undeserved suffering. But they should note that, had Smith's suffering been prevented, it would have been true that another moral agent no less good and just than Smith—a morally equivalent counterpart of Smith—endured the same preventable suffering that Smith endured in *w*. No one, including an Anselmian God, could ensure that there is no on-balance bad world at which a morally equivalent counterpart of Smith endures undeserved suffering. Call that the preventable evil thesis (PE).

PE. It is impossible that there should fail to be a bad world at which a good and just counterpart of Smith endures undeserved and preventable suffering. According to PE, no matter what anyone might have done in *w*, it would be true that there's a world in which a counterpart of Smith endures undeserved and preventable suffering.

So we know that either Smith endures undeserved and preventable suffering or some counterpart of Smith endures undeserved and preventable suffering. The suffering is equally bad whether Smith or some counterpart of Smith endures it. Suppose that an Anselmian perfect being exists in Smith's world. How could a perfect being allow Smith to endure the undeserved suffering? The theistic modal realist argues that, necessarily, the Anselmian God prevents Smith's suffering only if some counterpart of Smith endures undeserved suffering. An Anselmian God is not permitted to actualize the world in which Smith suffers undeservedly only if there is some moral reason why another counterpart of Smith ought to endure the suffering instead. But the relevant counterparts of Smith are no less good than Smith. So there is no moral reason why any of the relevant counterparts ought to endure the suffering rather than Smith.

The moral position of the Anselmian God is in perfect analogy to the moral position of a rescuer that can prevent each good and just person from drowning, but cannot prevent all good and just persons from drowning. **(p.7)** Call that a *Rescue Situation*. In Rescue Situations a rescuer is morally permitted to allow one person to drown if the loss of that life is necessary to preventing another person from drowning. Theistic modal realists urge that the Anselmian God is morally permitted to allow Smith to suffer undeservedly if doing so is necessary to preventing another equally good person from suffering.

VI. The Modal Realist's Formal Argument

It is definitive of Rescue Situations that some person R *can* save person S and *can* save person S´ but *cannot* save both S and S´. Since it is impossible for R to actualize a world in which both S and S´ are saved, R is permitted to save S at the necessary cost of not saving S´. It is difficult to find a thoughtful and well-informed person that is unwilling to accept that justification for saving one person and allowing another to drown.

But theistic modal realists have an analogous argument that justifies God in allowing Smith to suffer undeservedly.[16] The representation of the analogous Anselmian argument requires a domain that is suitably large. Counterpart theory provides a domain of quantification that includes every possible world and everything in every possible world. We include, among the existing objects, every possible object. We retain the familiar assumptions that properties are sets of possible objects and propositions are sets of possible worlds.

Theistic modal realists begin with the assumption that there is a set of worlds *W* in which God prevents the undeserved suffering that Smith actually endures. Each of these worlds includes a counterpart of Smith. Those counterparts are the representatives of Smith in those worlds. The relevant counterparts are moral equivalents of Smith. They are indeed Smith himself *according to* each of those worlds. Each of these counterparts is the person Smith would have been had Smith's suffering been prevented in some way.[17] Further, these worlds exhaust the possible ways in which God might have prevented the undeserved suffering that Smith actually endures.

The premises of the argument are set out in the language of counterpart theory. Let *Wy* state that *y* is a world, *Iyx* state *x* is in world *y*, *Cxa* state *x* is **(p.8)** a counterpart of *a*, *Mxa* state *x* is a moral equivalent of *a*, and *Ga* state that God prevent the suffering of *a*. Let Smith take the name *a* and suppose Smith exists in the actual world @. We assume that God necessarily exists and possesses all of the traditional divine attributes in every world.

The first premise in the argument states that Smith endures preventable suffering. Let *Sa* symbolize “Smith endures preventable suffering.” Tom Morris and Alvin Plantinga both reject the idea that any moral agent might endure pointless and preventable suffering.[18] The position of Plantinga and Morris leads to the conclusion that it is impossible that Smith suffer needlessly. But I side with the strong evidence that some actual moral agents do suffer needlessly.

1. $⍰(\exists x)(\forall z)(Izx \equiv Az) \;\&\; Sa_x)$

(1) is the translation of Sa—Smith (actually) endures preventable suffering—in the primitive notation of counterpart theory. In the logic of counterpart theory, all closed sentences are governed by Brouwer's Principle, which states⍰ $(p \supset ⍰\Diamond p)$. It follows from (1) and Brouwer's principle that it is *necessarily possible* that Smith endures preventable suffering.

2. $\Diamond Sa$

But that is just to say that it's necessary that some morally equal counterpart (or other) of Smith endures needless suffering. So, if God prevents the suffering of all of Smith's non-actual and morally equal counterparts, then God cannot also prevent the suffering of Smith.

Now, suppose for reductio that God prevents the suffering of every morally equal counterpart of Smith. In that case (3) is true.

3. $\forall y_1 \forall x_1\ (Wy_1 \;\&\; Iy_1x_1 \;\&\; \&\; Cx_1a \;\&\; Mx_1a \supset Gx_1)$

In modal talk, (3) states that God necessarily prevents Smith from suffering needlessly, or $\Box Ga$. It follows immediately that it is not possible that Smith endures preventable suffering, or $\sim\Diamond Sa$. But we know from premise (2) that it's necessarily possible that Smith endures preventable suffering. So premise **(p.9)** (3) is necessarily false, or $\sim\Diamond\Box Ga$. God simply cannot prevent the suffering of every morally equal counterpart of Smith.

But if (3) is necessarily false, then we know that (4) is necessarily true.

4. $\forall y_1 \forall x_1\ (Wy_1\ \&\ Iy_1x_1\ \&\ (y_1 \neq @)\ \&\ Cx_1a\ \&\ Mx_1a\ \&\ Gx_1) \supset {\sim}Ga$

Premise (4) states that if God prevents the unnecessary suffering of all of Smith's morally equal counterparts in every non-actual world, then God does not prevent the actual suffering of Smith. We take as the relevant counterparts of Smith all of those who are, like Smith, good and undeserving of the suffering they endure.

Suppose it's permissible that God prevents the unnecessary suffering of all of Smith's morally equal counterparts in every non-actual world.

5. $P(\forall y_1 \forall x_1\ (Wy_1\ \&\ Iy_1x_1\ \&\ (y_1 \neq @)\ \&\ Cx_1a\ \&\ Mx_1a\ \&\ Gx_1))$

It is difficult to see how (5) could be false. Certainly, it would seem, God is permitted to prevent the needless suffering of all of Smith's morally equal counterparts. But if premises (1)–(5) are true, then God is morally permitted not to prevent the unnecessary suffering of Smith.

6. $P{\sim}Ga$

The argument has the structure of the arguments offered in *Rescue Situations*. Recall that if a rescuer can save each of two persons, S and S´, but cannot save both S and S´, then he is permitted to save at least one. Similarly, if God can prevent the suffering of each of Smith's counterparts, but cannot prevent the suffering of all of Smith's counterparts, then God is permitted to prevent the suffering of all of Smith's counterparts except Smith himself.

The theistic modal realist assumes that it is possible that Smith suffers needlessly. That assumption seems reasonable. If Smith can suffer needlessly, then God simply cannot prevent every relevant counterpart of Smith from suffering in that way. The theistic modal realist does not assume that it is possible for God to *cause* suffering. That assumption is clearly unreasonable; God does not strongly actualize such states of affairs. So there is no world, the theistic modal realist should insist, in which a perfect being causes moral agents to suffer needlessly.

It might be objected that the Anselmian God ought not to be concerned about every moral agent that exists, but only about moral agents that *actually exist*. An Anselmian God that permits an actual moral agent to suffer undeservedly in order to prevent an existing, non-actual moral agent from suffering, does something morally wrong.

(p.10) But this objection is unmotivated. The suffering of other concrete universes is no less genuine than the suffering in our universe. We perhaps have special obligations to our world-mates. But certainly God's concern is with the multiverse as a whole.

The objection also undercuts the *Plenitude Problem* for theism. It entails that in non-actual worlds an essentially morally perfect being need not have a reason to allow a moral agent to suffer undeservedly and preventably. But of course it is false that in non-actual worlds an essentially morally perfect being may simply allow undeserved and preventable suffering. Unlike the rest of us, the Anselmian God is morally perfect in every world.

We should conclude that an essentially morally perfect being can permit undeserved and preventable suffering if it is the necessary cost of preventing equally bad suffering. In particular, an essentially morally perfect being is permitted to let Smith suffer undeservedly, if it is necessary to preventing his equally good, non-actual counterparts from suffering undeservedly.

VII. Less-Than-Best Problems Resolved

The theistic modal realist's solution to the *Less-than-Best Problem* makes the reasonable concession that our world is not the best possible world. But it should be obvious that the theistic modal realist will again respond that a perfect being is morally permitted not to make the lives of actual rational and sentient beings better. Certainly, a perfect being *can* improve the lives of every actual rational and sentient being, and certainly it would be a moral improvement if he did. But we know that, necessarily, if he or she does so, then some group or other of morally equivalent rational and sentient counterparts is such that their lives are not improved. It is impossible that a perfect being should improve the lives of every morally equivalent group of rational and sentient counterparts in every world. There is therefore no moral reason why a perfect being must improve the lives of all actual rational and sentient beings rather than improve the lives of their morally equivalent counterparts.

VIII. Other Worries for Theistic Modal Realism

An Anselmian God is a necessarily existing being, a being that exists in every possible world. Genuine modal realists are counterpart theorists. We all have counterparts in other worlds, but no one is identical to his or her **(p.11)** counterparts. In particular, God has counterparts in other worlds, but God is not identical to any of his counterparts. But then it appears that God—the greatest possible being in our world—actually exists, but does not necessarily exist.[19]

This worry is misplaced. It is true that God has counterparts in every world and it is true that each of those beings is world-bound.[20] But the Anselmian claim that God necessarily exists is also true. Indeed, that God has a counterpart in each possible world is just what it means to say that God necessarily exists. The worry that God does not necessarily exist simply reflects a global concern about the counterpart-theoretic equivalents of our familiar modal talk. It is not a particular concern for theistic modal realists.

It is true that God exists in a world *w* if and only if God is related to the parts of *w* in the right way. Among the best candidates for the right relation is standing in a spatio-temporal relation to the other parts of *w*. But other candidates are also reasonable. We assume that God is omnipresent, even if eternal. Since we can stand in a spatial relation to a God that is omnipresent, we are reasonably called world-mates with God.[21]

But suppose Anselmian perfect beings are essentially sole creators of everything that exists.[22] It is true that God essentially actualizes the world, but it is not true that God created everything unrestrictedly. God did not create all of the possible objects that exist, for instance; he possibly created the possible objects that exist. So there are objects that the Anselmian God did not create.

I don't find this is an especially large obstacle for theistic modal realists. The Anselmian God essentially actualizes the world. This is to say that the Anselmian God essentially weakly actualizes every contingent state of affairs that obtains. He creates the inhabitants of every world, so in some sense he creates necessarily existing objects as well—objects that have counterparts **(p.12)** in every world. Indeed, since he creates all actual objects and all actual objects necessarily exist, God creates large numbers of necessarily existing beings.[23] But there is no conceptual problem in claiming both that the Anselmian God is essentially sole creator and that there exist some non-actual objects that he did not actually create.

The Anselmian God is also essentially omniscient. It is typical to analyze omniscience in terms of propositional belief and knowledge. The natural proposal is that an omniscient being is one that believes all and only true propositions and knows every proposition he believes. But that can't be quite right since there are true propositions that are *not definitely* true and there are definitely true propositions that are not definitely, definitely true, and so on.[24] The problem is more pronounced for modal realists who countenance transworld objects. It will be indeterminate whether an object is actual when many, but not all, of its parts are actual. For all such transworld objects x—and there will be infinitely many for the unrestricted compositionalist—it is true, but not definitely true that x is actual. An Anselmian God would not know that any of those objects are actual objects.

This again is not a problem for theistic modal realism. First, theistic modal realists need not endorse unrestricted composition. Second, even if they do, it is not a large concession that an omniscient being does not know any proposition that is indefinite at any order of vagueness. For any object x, if x is on any borderline of the property F, then an omniscient being does not know that Fx is true.[25] It is perfectly reasonable to hold that omniscient beings know all true propositions, modal and non-modal, that are neither higher nor lower order vague.

There is also a worry about true, necessarily false propositions. Propositions stating that a transworld object exists can be true, though false in every possible world. Since the object does not exist (restrictedly) in any world, it is not true that “object O exists” is possibly true. Yet, it is true that there exists (unrestrictedly) the transworld object O. Does God know that the object unrestrictedly exists?[26] I don't think theistic modal realists have any problem in answering affirmatively, God does know some true propositions that are necessarily false.

(p.13) Theistic modal realism can also handle the long-standing objection that since Anselmian perfect beings are *impeccable* they cannot also be essentially omnipotent.[27] An omnipotent being would be able to perform any possible action including immoral actions.

There are several responses to the impeccability objection including arguments from necessarily unmanifested powers.[28] None of these is especially plausible. But theistic modal realists needn't concede that an Anselmian perfect being cannot perform an immoral action. Indeed, there is nothing inconsistent in the claim that an impeccable being *can* perform an immoral action.

Compare the possibility of a time-traveler killing his own grandfather. To say that a time-traveler can kill his own grandfather is just to say that his doing so is compossible with certain facts. According to counterpart theorists, determining which compossible facts are relevant is a contextual matter. To suppose that Tim were to kill his own grandfather is to suppose that Tim does so in some set of relevantly similar worlds. The similar worlds are certainly not ones in which a contradiction is true.

Let *A* be the immoral action of harming Smith. What would have to be the case were an Anselmian perfect being to perform *A*? It certainly would not be the case that God is both impeccable and not impeccable. No contradiction would follow. It would not be the case that God fails to be perfectly good, since God is essentially perfectly good. So, were God to perform *A*, it *would not* be an instance of immorally harming Smith. Perhaps it would be an instance of not harming Smith at all, or an instance of justifiably harming Smith. In any case, it is true that *A* is an instance of immorally harming Smith and it is possible for an impeccable being to perform *A*.[29] Theistic modal realists therefore have a solution to the long-standing objection to Anselmian accounts of omnipotence that allows perfect beings to perform immoral actions.

IX. Conclusions

Theistic modal realists have the resources to explain how God might allow a moral agent to suffer needlessly. God may allow Smith to suffer needlessly, for instance, if God prevents the suffering of every morally equivalent counterpart of Smith. God's situation is what I have called a *Rescue Situation*. The very same reasons that justify a rescuer in saving some, but not all, **(p.14)** drowning persons, justify God in preventing the suffering of some, but not all, counterparts of Smith.

Theistic modal realists also have the resources to explain how God might actualize a less-than-best world. It is necessarily true that some world *w* is less than the best possible world. So, it is necessary that, if God improves the lives of every sentient member *S* of *w*, then he fails to improve the lives of the counterparts of *S* in some world *w*´. But there is no moral reason why the sentient members of *w*´ should have lives that are less good than the sentient members of *w*. So God has no moral reason to improve the lives of the sentient members of *w*.

We considered, finally, some of the traditional attributes of the Anselmian God in the context of theistic modal realism. We found that theistic modal realists can offer some intriguing accounts of the attributes of necessary existence, essential omniscience, essential omnipotence, essential perfect goodness, and being essential sole creator. It seems reasonable to conclude that theistic modal realism offers a fascinating new avenue for illuminating Anselmian perfection.

References

Bibliography references:

Almeida, Michael J. (2008), *The Metaphysics of Perfect Beings* (London: Routledge), section 3.10.

Cameron, Ross (2009), “God Exists at Every Modal Realist World: Response to Sheehy,” *Religious Studies* 45: 95–100.

—— (2009), “What's Metaphysical about Metaphysical Necessity?” *Philosophy and Phenomenological Research*, 1–16.

Divers, John (2002), *Possible Worlds* (New York: Routledge), 46 ff.

Funkhouser, Eric (2006), “On Privileging God's Moral Goodness,” *Faith and Philosophy* 24(3): 409–22.

Garcia, Laura (1984), “A Response to the Modal Problem of Evil,” *Faith and Philosophy* 1: 378–88.

Guleserian, Theodore (1983), “God and Possible Worlds: The Modal Problem of Evil,” *Noûs* 17: 221–38.

Hawthorne, John (2005), “Vagueness and the Mind of God,” *Philosophical Studies* 122: 1–25.

Heller, Mark (2003), “The Immorality of Modal Realism, Or: How I Learned to Stop Worrying and Let the Children Drown,” *Philosophical Studies* 114: 1–22.

Hudson, Hud (1997), “Brute Facts,” *Australasian Journal of Philosophy*, 77–82.

Lewis, David (1983), “Counterpart Theory and Quantified Modal Logic,” *Philosophical Papers* I (Oxford: Oxford University Press), 28 ff.

(p.15) —— (1986), *On the Plurality of Worlds* (Oxford: Blackwell Publishers), 86.

Morris, Thomas V. (1987), “The Necessity of God's Goodness,” in his *Anselmian Explorations: Essays in Philosophical Theology* (Indiana: University of Notre Dame Press), 42–69.

Nolan, Daniel (2005), *David Lewis* (Quebec, Canada: McGill-Queen's Press), 55 ff.

Quinn, Phil (1982), “God, Moral Perfection and Possible Worlds,” Sontag and Bryant, 212.

Rowe, William (1996), “The Problem of Evil and Some Varieties of Atheism,” collected in Daniel Howard-Snyder (ed.), *The Evidential Argument from Evil* (Indianapolis: Indiana University Press), 1–11.

Senor, Tom (2006), “God's Goodness Needs No Privilege: A Reply to Funkhouser,” *Faith and Philosophy* (24)3: 423–31.

Sheehy, Paul (2006), “Theism and Modal Realism,” *Religious Studies* 42: 315–28.

Sontag, Frederick, and Bryant, M. Darrol (1982) (eds.), *God: The Contemporary Discussion* (New York: Rose of Sharon Press), 212.

Notes:

(*) My thanks to Ross Cameron, John Greco, Jon Kvanvig, Chris Menzel, and Stuart Rosenbaum for discussion, comment, and criticism of various issues in this paper. I thank also two referees from Oxford and the participants in the *Third Annual Baylor Philosophy of Religion Conference*.

(1) Quinn (1982: 212).

(2) Phil Quinn is not especially clear on this issue. He defines an actualizable world as a world that an omnipotent being could actualize. Since that definition is nearly trivial, it remains unclear whether an omnipotent being could actualize every logically possible world. On the other hand, he is explicit in wanting not to decide the issue either way. See ibid. 205 ff).

(3) See Nolan (2005: 55 ff.).

(4) See Heller (2003: 1–22).

(5) See Divers (2002: 46 ff.).

(6) Anselmian eternalism is assumed here to be compatible with God's omnipresence. Since we stand in a spatio-temporal relation (or a close analogue of a spatio-temporal relation) to a God that is omnipresent, even if that being is atemporal, we are to that extent world-mates with God.

(7) Compare note 19 below. It is sometimes complained that Anselmianism is incompatible with Lewisian modal realism (plus counterpart theory) since the latter entails that God is world-bound. As far as I can tell, the objection confuses being world-bound with not necessarily existing. The Anselmian view of God does not entail that God is not world-bound. Rather Anselmianism entails that God is *necessarily existing*, and that is perfectly compatible with Lewisian modal realism together with counterpart theory.

(8) See Lewis (1986: 86).

(9) Ibid. (87–8).

(10) It is not clear that the *Principle of Plenitude* entails that there would be an infinite number of worlds or, as Peter van Inwagen has objected, that there would be more than seventeen possible worlds. It is an important problem, but I do not address it here. See Lewis (1986: 86).

(11) See Guleserian (1983: 221–38).

(12) Ibid. (224).

(13) Of course, strictly, each world is a real concrete universe that a perfect being has actualized *at that world*.

(14) See Morris (1987: 42–69).

(15) See Rowe (1996: 1–11): “An omniscient, wholly good being would prevent the occurrence of any intense evil it could, unless it could not do so without thereby losing some greater good or permitting some evil equally bad or worse.”

(16) Compare Garcia (1984: 378–88).

(17) See Lewis (1983: 28 ff.).

(18) This needs qualification. Plantinga of course allows that God might permit moral agents to act freely in ways that are harmful to other moral agents. The classic source is his *The Nature of Necessity* (Oxford: Clarendon Press, 1974). But the value in exercising freedom, or the value of beings that can exercise freedom, is supposed to counterbalance those moral evils. The account is of course much more complex and subtle. Plantinga's account is plausible in its explanation of how it would be worse overall were God to prevent moral evil. In the cases I'm discussing, it would not be worse overall in Smith's world were God to prevent Smith from suffering. It would indeed be better overall in that world.

(19) Compare note 7 above. The objection was raised by Stuart Rosenbaum and Jon Kvanvig in discussion, but it is a familiar one. It is also raised in Sheehy (2006: 315–28). For a cogent response to many of the worries raised in Sheehy's paper, see Cameron (2009: 95–100). Cameron proposes several ways in which a modal realist might accommodate a necessarily existing being. As far as I can see, Sheehy is concerned with a pseudo-problem based on a misunderstanding of the reductive project of counterpart theory and genuine modal realism. Genuine modal realism offers a way to say everything we have wanted to say about the necessary existence of the Anselmian God.

(20) I sidestep concerns about the number of counterparts God might have in any particular world.

(21) For other possibilities, see Lewis (1986: 73 ff.).

(22) Ross Cameron suggested in correspondence that this might be a problem for the Anselmian God in my conception.

(23) See Lewis (1983: 26–54). Lewis notes, p. 31 ff., that $\forall x \Box (\exists y)(x = y)$ come out true, given his translation schemes. Of course, counterpart theory is a *theory*, so such schemes are certainly subject to modification.

(24) For a nice discussion of related points, see Hawthorne (2005: 1–25). For problems with Hawthorne's proposal, see Almeida (2008: 3.10).

(25) See Almeida (2008: 3.9–3.11).

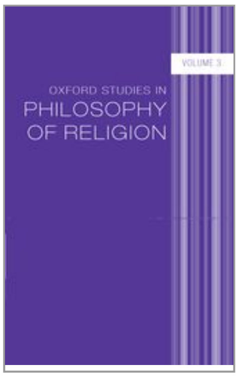
---

(26) Compare, for instance, Hudson (1997: 77–82).

(27) See Funkhouser (2006: 409–22).

(28) Compare, for instance, Senor (2006: 423–31).

(29) Compare Cameron (2009: 1–16).



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

The Argument from Miracles *

Daniel Bonevac

DOI:10.1093/acprof:oso/9780199603213.003.0002

Abstract and Keywords

This chapter defends the argument from miracles, an argument for God's existence that is primarily historical. It does not depend on any particular definition or description of God. It yields, if successful, de re knowledge of God. And it expresses an important reason why many believers believe in God. The chapter casts the argument in a general form, brings out its recursive character, shows that similar considerations apply to base and recursive portions, and demonstrates that the argument becomes quite strong when applied to events with multiple witnesses and events in series.

Keywords: miracles, historical argument, existence, witnesses, series

Philosophers and theologians have advanced a variety of arguments for God's existence. The arguments vary along a number of dimensions. Some are a priori; some are a posteriori. Some depend on a particular definition of God; others are more general. But they have features in common. They are independent of any particular religious tradition, depending on nothing more than a generally classical conception of God.[1] They attempt to relate a potential believer to God by way of a definition or description. They thus appear, if successful, to yield *de dicto* rather than *de re* knowledge of God. And they bear little relation to the reasons for which most believers believe.

My aim, in this paper, is to revive an argument for God's existence that does not share these features. It is primarily historical; it does not depend on any particular definition or description of God. It yields, if successful, *de re* knowledge of God within the context and history of a particular religious **(p.17)** tradition. And it does express an important reason why many believers believe in God. Indeed, it formed the central argument for Christianity both in ancient times and in Enlightenment debates about the justification for religious belief.

The argument I will defend is the argument from miracles, also known as the historical argument. It has been receiving increased attention. Bayesian methods have revealed weaknesses in David Hume's critique of the argument, and the so-called higher criticism that cast doubt on the historicity of the Bible increasingly appears methodologically unsound.[2] My goal in this paper is to cast the argument from miracles in a general form, bring out its recursive character, show that similar considerations apply to base and recursive portions, and demonstrate, in particular, that the argument becomes quite strong when applied to events in series.

I. The Argument

The primary, and perhaps the only, argument for the existence of God in the Old and New Testaments and the early Church Fathers is the argument from miracles. In his first letter to the Corinthians, for example, Paul writes:

> First and foremost, I handed on to you the facts which had been imparted to me: that Christ died for our sins, in accordance with the scriptures; that he was buried; that he was raised to life on the third day, according to the scriptures; and that he appeared to Cephas, and afterwards to the Twelve. Then he appeared to over five hundred of our brothers at once, most of whom are still alive, although some have died. Then he appeared to James, and afterwards to all the apostles. In the end he appeared even to me. (I Corinthians 15:3–8)

Paul takes this as the “first and foremost” aspect of the Gospel he preached in Corinth, the set of facts on which he bases his faith. The Easter message that Christ rose from the dead, Paul holds, is absolutely essential: “And if **(p.18)** Christ was not raised, then our gospel is null and void, and so is your faith; and we turn out to be lying witnesses for God” (I Corinthians 15:14–15). But Paul assures us that Christ *was* raised to life; there is the solid testimony of a multitude of eyewitnesses to Christ's resurrection. Christian belief, Paul maintains, rests on firm empirical evidence.

Philosophers, however, have not given such empirical evidence much respect for the past quarter-millennium. Hume, in particular, is generally thought to have dealt a knockout blow to attempts to base religious faith on the occurrence of miracles. Intellectual historian John Herman Randall, Jr., representatively takes Hume to have delivered a “coup de grace,” for he:

> proved so conclusively that intelligent men have rarely questioned it since, that a miracle, in the sense of a supernatural event as a sign of the divinity of its worker, cannot possibly be established.[3]

Together with the assaults on natural theology launched in Hume's *Dialogues Concerning Natural Religion* and Kant's *Critique of Pure Reason*, Hume's punch devastated not just a belief in miracles but any attempt to rest religious faith on a rational or objective foundation.

But Hume does not argue that miracles are impossible. Indeed, his account of natural laws as nothing more than regularities in nature precludes that strategy. Instead, Hume argues that it would never be rational to accept a miracle report, that is, a bit of testimony that a miracle had occurred. It is far more likely, he contends, that the Apostles and the five hundred were mistaken or lying than that anyone was actually raised from the dead.

Though there are some critical voices, most published discussions of Hume's critique find it compelling. Here are two typical conclusions, on opposite ends of the twentieth century. Leslie Stephen: “The statement that a man has been raised from the dead would prove that its author was a liar.” Howard Sobel: “If anyone were to tell you that a man had died and come back to life you had better not believe him.”[4]

I shall argue that these conclusions are far too sweeping. Hume's argument is compelling at best in establishing what Roy Sorensen calls “case-by-case skepticism” about miracles of a certain kind.[5] That is, Hume shows that it would never be rational to accept a single miracle report in isolation, where the miracle in question involves a violation of natural law. That, however, does nothing to undermine the miracles that form part of the foundation of Christian belief.

(p.19) I shall formulate the argument from miracles, as a first pass, as follows:

1. There are kinds of possible circumstances and events the best explanations for which invoke supernatural agency.
2. Some circumstances and events of those kinds have actually occurred.
3. Therefore, there is a supernatural agent.

This argument raises some obvious questions. What kinds of circumstances and events? When is agency of any kind the best explanation for an event? Can supernatural agency ever be the best explanation for something? Why should we even believe that supernatural agency is possible? What is it to be supernatural? Even if some events are best explained by appeal to supernatural agency, why think that the same supernatural agent is active in each case? Why think that the supernatural agent in each case is God? I cannot answer all these questions here. I shall be content to suggest answers to some of them, taking for granted our common-sense intuitions about attributions of agency and even of supernatural agency.

The second premise, that circumstances and events of the relevant kind have actually occurred, raises further issues of its own.[6] Contemporary Judeo-Christian believers base their beliefs on occurrences in the relatively distant past. That allows us to divide the question, and move to the forefront the epistemological questions that concerned Hume. Under what conditions is it rational to believe that events of the relevant kind have actually occurred? To answer that question, we need to know under what conditions it would be rational to give credence to testimony that such events have occurred—the question upon which most of the literature concentrates—and also under what conditions it would be rational to understand one's *current* experience as being of the relevant kind, of a kind best understood in terms of supernatural agency. There is, for example, not only the issue of whether *we* should accept the account of Moses's encounter with the burning bush (Exodus 3:2) but whether Moses *himself* should have understood his experience in terms of supernatural agency, and even whether Moses should have believed that a bush was burning without being consumed, rather than concluding that he was suffering some sort of illusion or hallucination.

The argument from miracles, like any historical argument, thus has a kind of recursive structure. There is a base case, an argument that certain kinds of events have a certain property. There is also an inductive step, an **(p.20)** argument that certain kinds of events have that property if other preceding events have had it as well. The property I am interested in here is that of supporting rational credence in supernatural agency. So, the argument will proceed by identifying initial cases directly supporting such rational credence and then specifying conditions under which rational credence may be transmitted through testimony. By analogy to Robert Nozick's historical theory of justice, we might think of these as conditions of acquisition and transfer of rational credence.[7] The idea, then, is that one's epistemic situation supports rational belief in supernatural agency if and only if that situation is an initial acquisition of rational credence—that is, includes experience of an event that directly supports such belief—or is a transfer of rational belief in supernatural agency from someone who holds a corresponding belief rationally.[8]

I want to caution against a possible misunderstanding, and, in the process, complicate this picture somewhat. To say that some events might support rational credence directly is not to say that they support it immediately or in isolation. It might be, for example, that an initial experience of the relevant kind would not support rational credence, but that later, similar experiences would. As Hume stresses in his critique of causation, one instance of an event of kind A followed by an event of kind B does not tend to produce belief in a necessary connection between A and B. Repeated instances do. Similarly, one encounter with a bush that burns but is not consumed might not support rational belief in supernatural agency, but a series of events beginning with that one might.[9] We might follow the literature on miracles here and speak glibly of “events or series of events,” “miracles or series of miracles,” etcetera. But this obscures the issues considerably by ignoring the context-sensitivity of the relevant epistemic notions. Whether or not an event supports rational credence in something is plainly context-sensitive; it depends **(p.21)** not on that event in isolation but on previous events. We'll investigate some aspects of this context-sensitivity below. For now, let's simply say that an event may support rational credence in supernatural agency by being a contributing term in a series of events that support such credence.

II. Miracles

The argument from miracles begins by asserting that there are kinds of possible events the best explanation for which would be supernatural agency. What kinds of events are those? When, if ever, could we best explain an event by invoking supernatural agency? Traditionally, events of this kind are known as miracles. What is a miracle? All the argument requires is that miracles be such that the best explanations for them invoke supernatural agency. Adopting that as the definition of miracles for the purposes of the argument allows us to simplify it:

1. *Definition*. Miracles are events the best explanation for which would invoke supernatural agency.
2. *Possibility*. Certain kinds of possible events are miracles.
3. *Actuality*. Events of those kinds have actually occurred.
4. Therefore, there is a supernatural agent.

This definition of miracles is admittedly epistemological. One could take that at face value, or see it as a placeholder for a more basic metaphysical definition. Here I shall treat its metaphysical neutrality as a virtue, for it is consistent with various ways of accounting for agency in general and supernatural agency in particular. The definition is also weaker than most in the literature. It is worth pausing to explore the contrast.

Hume defines miracles as violations of the laws of nature:

> A miracle is a violation of the laws of nature; and as a firm and unalterable experience has established these laws, the proof against a miracle, from the very nature of the fact, is as entire as any argument from experience can possibly be imagined.[10]

But this seems neither necessary nor sufficient for miracles in the traditional sense. Certainly it is neither necessary nor sufficient for having supernatural agency as a best explanation. It is not necessary, for many Biblical miracles involve no violation of laws of nature. If we look, for example, at the plagues of Egypt, we find several—plagues of flies (Exodus 8:20–5), livestock illnesses **(p. 22)** (9:1–7), boils (9:8–12), hail (9:22–6), and locusts (10:12–15)—that require no suspension of the laws of physics. Neither do some of the miracles of Elijah or Elisha—the bear attack (2 Kings 2:23–5), for example, or neutralizing the poisoned stew (2 Kings 4:38–41). Neither is it sufficient. The laws of planetary motion are well understood. Suppose astronomers were to discover a slight deviation in a single orbit of Neptune, say, for which they could give no account. They would of course hypothesize a force causing the deviation. But suppose, after extensive investigation, that no cause could be found. It might be rational to conclude, after enough investigation and theoretical reflection, that there had been a violation of natural law. But barring any effect that could lead us to think of the planet's deviation as fulfilling some purpose, we would not tend to count it as a miracle. Thinking of miracles as violations of natural law cannot explain the relevance of purpose here. But thinking of them as events the best explanation for which invokes supernatural agency does, for agency is purposive action.[11]

There is another problem with Hume's definition. Some of the laws of our best current scientific theories are statistical. It is not clear what it would mean for a single event to violate such laws. According to statistical mechanics, it is possible that all the air in this room gathers in one corner—it is just extremely unlikely. Suppose it were to happen, causing a group of criminals to collapse from lack of oxygen. Would that count as a violation of natural law? It is not inconsistent with the laws of nature, but they assign it such low probability that it might be reasonable to take the event as a sign of supernatural agency. A similar point could be made about quantum tunneling. Extremely low probability events do not violate laws, but might or might not seem to point to supernatural agency, depending on whether we can construe them as purposive.

Timothy and Lydia McGrew define miracles as events that would not have occurred in the natural order of things. This is very close to my definition, for, if an event would not have occurred in the natural order of things, one might think, the best explanation for it must not be completely natural.[12] Their definition is in one way broader than mine, however, for events **(p.23)** that lack fully naturalistic explanations might best be explained in terms of supernatural agency, but they might simply be inexplicable. Or, the best explanations for them might not invoke agency, even if they invoke something supernatural. Let's return to the minor and evidently purposeless deviation in Neptune's orbit. If we can find no naturalistic explanation, we might conclude that the event could not be explained at all. We might, alternatively, hypothesize some sort of impersonal force having a source outside the natural order as usually understood. The McGrews' definition would count the deviation miraculous. My definition would not.

The McGrews' definition may be narrower than mine in another way, depending on what they take the scope of the natural order to be, for the natural order, on their view, is by definition unmiraculous. One could see the cosmological argument, the anthropic argument, and the argument from design, however, as arguments that the best explanation for the natural order itself invokes supernatural agency. If those arguments were successful, then, my definition, for better or worse, would account the existence of the natural world itself as miraculous. I view that as a strength; others may disagree.[13] In any case, nothing in my argument here will depend on this feature of the definition.

III. Hume's Argument

Redefining miracles in my way or in the McGrews' way might seem to lead to a quick refutation of Hume's critique. Hume's argument addresses the transfer principle, and shows, at best, that it would never be rational to accept a report of a violation of a law of nature. If some miracles involve no violation of natural law, this is consistent with rational credence of some miracle reports.

Such a response, however, concedes far too much. Most Biblical miracles, and especially most of the miracles of Jesus, do violate natural law. It is hard to give any naturalistic explanation of the Passover, of Elisha's raising the son of the Shunammite woman from the dead (2 Kings 4:18–37), or, especially, of Jesus's resurrection. So, we must face Hume's argument squarely. If it succeeds, it not only undercuts any justification for faith based on miracles, **(p.24)** but implies that faith in the Biblical story is irrational. We might still affirm with Paul that “Divine folly is wiser than the wisdom of man” (I Corinthians 1:25), but we would have to agree with Hume that faith is folly.

Let me begin by outlining Hume's argument, which, he believes, “must at least silence the most arrogant bigotry and superstition, and free us from their impertinent solicitations” by being “an everlasting check to all kinds of superstitious delusion.”[14] Hume is not arguing that miracles are impossible. It has been argued that miracles are violations of laws of nature; but laws of nature, by definition, are universally valid; therefore, miracles cannot occur.[15] But that is not Hume's point. All reasoning concerning matters of fact, Hume insists, rests on experience, and is properly understood as a matter of probability. “A wise man, therefore,” says Hume, “proportions his belief to the evidence.”[16] We have no basis for saying that miracles are impossible, therefore; we can say only that they are very unlikely. The evidence is overwhelmingly against them.

Hume argues that miracles are so unlikely that we never have good reason to believe in them. Our only evidence for miracles—certainly, for the miracles of the Old and New Testaments—is the testimony of eyewitnesses reported in the Bible. But we have ample experience of people testifying falsely. We must assess the credibility of these reports in the way we assess the credibility of any testimony.

Now, at first glance, this would seem to lend credence to Biblical miracles, for we have reports of these miracles occurring; we have no corresponding eyewitness reports that contradict them. We have no records of men claiming to have impregnated Mary, for example, or claiming to have seen Christ's body in the tomb a week after Easter. But the story is not so simple, for miraculous events are extremely—Hume might even say incredibly—unlikely. And the credibility of testimony depends in part on the likelihood of what is being reported:

> Suppose, for instance, that the fact, which the testimony endeavours to establish, partakes of the extraordinary and the marvellous; in that case, the evidence, resulting from the testimony, admits of a diminution, greater or less, in proportion as the fact is more or less unusual…I should not believe such a story were it told me by Cato, was a proverbial saying in Rome, even during the lifetime of that philosophical patriot. The incredibility of a fact, it was allowed, might invalidate so great an authority.[17]

(p.25) Thus, I believe my brother if he tells me that his daughters have colds, but not if he tells me that a flying saucer followed him to work. Some things are so unlikely that we consider them incredible. That applies, Hume argues, to the Resurrection and other purported miracles. Hume holds that miracles are so unlikely that no amount of testimony would convince a rational person of their occurrence. He concludes:

> Upon the whole, then, it appears, that no testimony for any kind of miracle has ever amounted to a probability, much less to a proof…[18]

IV. Transfer: Multiple Witnesses

Hume's argument, however, fails, even for miracles that violate the laws of nature. I am arguing that the central miracles of the Judeo-Christian tradition survive Hume's assault. Hume fails to show that belief in them is irrational. I am not arguing that belief in them is rationally compelled; that depends crucially on empirical matters concerning the likelihood of certain kinds of events, the reliability of witnesses, and the accuracy of the historical record. I am not, in short, concerned to argue in favor of the actuality premise, the premise that miracles have actually occurred. That is a matter, from my point of view, of detailed historical investigation and debate. But I am arguing that we cannot rule the argument from miracles out without such historical investigation and debate. Belief in Biblical miracles is not a priori irrational. Faith in miracles is not necessarily folly.

I will be concentrating on two flaws that rescue the central miracles of the Christian tradition from Hume's assault.

First, Paul is right to stress the number of eyewitnesses. Hume shows that it would not be rational to take a single witness as more credible than an established law of nature. But it might be rational to accept the testimony of several witnesses that a law of nature had been violated. The central miracles of the Judeo-Christian tradition are attested by multiple witnesses; Hume's argument leaves them untouched.

Most of the central miracles on which Christian faith depends are not the mystical experiences of a guru or divine revelations to a single prophet but publicly observable occurrences witnessed by many people. Jesus, to be sure, seeks to preserve “the Messianic secret,” limiting the number of people who witness certain miracles and know he is the Messiah. He allows only Peter, **(p. 26)** James, and John to accompany him when he revives the twelve-year-old girl, Jairus' daughter (according to Mark 5:35–43 and Luke 8:49–56; but cf. Matthew 9:18–19, 23–5). He asks the blind man from Bethesda not to tell anyone in the village that Jesus restored his sight (Mark 8:22–6; cf. Matthew 9:27–31). He gives the disciples strict orders not to tell anyone who he really is (Mark 8:30; Matthew 16:20). But even the most private miracles, such as the Transfiguration, are witnessed by several disciples.

To understand the effect this has on the credibility of miracle reports, we must put Hume's argument in more formal probabilistic terms. A variety of contemporary writers (Richard Swinburne, Howard Sobel, David Owen, George Schlesinger, Peter Millican, Philip Dawid and Donald Gillies, among others),[19] have interpreted Hume's argument in modern probabilistic, and specifically Bayesian, terms. Hume's argument is in various ways un-Bayesian; he distinguishes “proof” from “probability” in a way that makes no Bayesian (or for that matter Humean!) sense, and speaks of “subtracting” and “demolishing” probabilities. Nevertheless, a Bayesian framework allows a useful and formally precise reconstruction of something very close to Hume's argument. And no alternative probabilistic framework makes any better sense of Hume's language. To assess Hume's success, therefore, let's recast his argument in Bayesian terms.

Bayes's theorem allows us to calculate the conditional probability of an event in a context (K) from various other conditional probabilities in that context. In our setting—assessing the credibility of testimony concerning miracles—the conditional probability we are interested in is credibility, the probability that a miracle occurred given the testimony that it occurred ($P[M/(T\&K)]$). Bayes's formula equates that with a ratio involving the prior probability of such a miracle ($P[M/K]$) and the reliability of the witness: the likelihood of the witness saying that the miracle occurred, when it did occur ($P[T/(M\&K)]$), together with the likelihood of the witness saying that the miracle occurred, when it did not occur ($P[T/(-M\&K)]$):

$$P[M/(T\&K)] = \frac{P[T/(M\&K)]P[M/K]}{P[T/(M\&K)]P[M/K] + P[T/(-M\&K)]P[-M/K]}$$

The more reliable the witness, the greater the credibility of the testimony. But also, the more unlikely the event to which the witness is testifying, the smaller the credibility of the testimony.

(p.27) Before considering numerical values, let's simplify a bit. Let c_K, credibility in context K, represent $P[M/(T\&K)]$, and let p_K, the prior probability of a miracle occurring in K, represent $P[M/K]$. Let's assign it the value 10^{-m}. Let l_K, the probability that the witness giving a miracle report is lying in K, represent $P[T/(-M\&K)]$, and assume it has the value 10^{-r}. Then, assuming that the probability that someone witnessing a miracle will report it is relatively high, and that l_K is significantly greater than p_K, Bayes's formula allows us to approximate the credibility of a witness in context K as follows:[20]

$$c_K \approx \frac{p_k}{l_k} = 10^{r-m}.$$

Now, to get actual probabilities out of Bayes's theorem, we need to have values for the prior probability of the miraculous event occurring and values for the reliability of the witness or witnesses. Assessing these is of course immensely difficult. But let's make a rough estimate for a report claiming someone to have been raised from the dead.

First, we need to estimate the prior probability of such an event. The Bible contains several such reports, but their veracity is in question. Since the beginning of time there have been, within an order of magnitude or so, about ten billion human beings on the planet. And there have been only a few scattered reports of resurrections, whose credibility is in question. So, let's estimate the probability of resurrection, given the available evidence, at 1 in 10 billion: 10^{-10}.[21]

The reliability of witnesses is perhaps easier to estimate. People are generally reliable, especially on matters such as whether someone is walking on or through the water, whether someone is alive or dead, etcetera. Indeed, as Donald Davidson has argued, the possibility of linguistic communication depends on such reliability. In the case of a miracle report, we must estimate the probability that someone, a disciple of Jesus, say, will report a miracle if it occurs. Presumably the probability is very high, though it is not 1, as Peter's denial of Jesus illustrates. So, let's estimate the probability, cautiously, at .99.[22] What about the probability that someone will report a miracle even if one does not occur? David Owen and others have assumed, given the values we've estimated so far, that this will be unlikely, having probability **(p.28)** .01. Hume clearly thinks it is higher; disciples having a tendency to inflate the reputation of their leader. Still, very few spiritual leaders have been alleged to have the ability to raise people from the dead. (No such reports are associated with Confucius, Laozi, the Buddha, Zoroaster, or Mohammed, for example.) So, let's estimate this, cautiously, at .1.

Now, given these estimates, Bayes's theorem tells us that the probability of someone's being raised from the dead, given testimony to such an event, is approximately 10^{-9}: one in a billion. Hume appears to be vindicated. The probability we should rationally assign to someone's being raised from the dead, even given testimony that it has occurred, is very low. Even if we abandon our cautious estimates above, raising the witness's reliability to .999 and lowering the likelihood of a false report to .01, the odds of the report's being correct are approximately 10^{-8}, one in a hundred million. Hume is right that “no testimony is sufficient to establish a miracle, unless the testimony be of such a kind, that its falsehood would be more miraculous, than the fact, which it endeavours to establish” **(90)**—so long as the miracle in question is isolated, violates a law of nature (or at least has an extremely low prior probability), and is attested by a single witness.

But now, suppose that we have not one witness but several. As John Earman and Rodney Holder have observed, having multiple witnesses changes the outcome of our assessment of miracle reports dramatically.[23] Oddly, few other philosophers have thought the number of witnesses makes any difference. Dawid, Gillies, and Sobel, for instance, speak simply of "a witness or group of witnesses." Yet an analogy to law should suggest that this is absurd. It matters how many independent witnesses testify similarly. One witness who identifies the perpetrator has some effect on the probability of guilt or innocence; a dozen who independently do so have a much more powerful effect.

If we were to take Hume's argument as showing that testimony can never establish the likelihood of a miracle, as he wants us to, it would prove too much. Hume's argument depends solely on the thought that miracles are extremely unlikely. So, his argument should apply to anything that has a very low prior probability. It thus, if successful, will imply that we can never be rationally justified in believing that an extremely unlikely event has actually occurred. But that is outrageous.

Consider a situation that might be represented by similar calculations: a case of medical diagnosis.[24] Suppose that a highly reliable test diagnoses you as having an exceptionally rare disease. Say that the reliability of the test **(p.29)** is .999; it is wrong in only one case in a thousand. And suppose the disease is very rare, afflicting only one person in a million. What is the probability that you actually have the disease? According to Bayes's theorem, only about $10^{3-6} = 10^{-3}$, that is, about one in a thousand! Although the test is right 999 times out of a thousand, its positive result in your case will be a false positive 999 times out of a thousand.

This result is surprising. But think of how the test might function applied to all the roughly 300 million residents of the United States. About 300 would have the disease, and the test would accurately give a positive result for (nearly) all of them. But 299,999,700 people would not have the disease, and the test, wrong only one time in a thousand, would nevertheless produce about 300,000 false positives. So, the test, applied to the population of the U.S., would come up positive 300,300 times, and be right in only 300 of them. We tend to ignore base rates (that is, low priori probabilities) in our thinking, something some psychologists have dubbed a "cognitive illusion." So, one test, even if it is highly reliable, is not very good evidence that any particular person has a rare disease.

But it would be absurd to conclude from this that we can never have good reason to believe that any particular person has a rare disease. True, any single test, taken by itself, is poor evidence. But, faced with a positive result, what might we do? We might repeat the same test. We might administer additional tests. We might look for symptoms. In short, we might gather additional evidence.

Analogously, faced with a miracle report, we ought rationally to gather additional evidence. Just as we might seek additional tests, we might, for example, seek testimony of additional and independent witnesses. Suppose we have *n* independent witnesses, all of equal reliability. Then Bayes's theorem tells us that the credibility of their reports, taken together, is:

$$P[M/(T\&K)] = \frac{P[T/(M\&K)]^n P[M/K]}{P[T/(M\&K)]^n P[M/K] + P[T/(-M\&K)]^n P[-M/K]}$$

Approximating as before, we get something that breaks down as *nr* approaches *m:*

$$c_k \approx \frac{p_k}{l_k^n} = 10^{nr-m}$$

Let's apply this to the medical diagnosis case. We have 300,000 false positives and only 300 true positives. Suppose we apply a second medical test **(p.30)** the reliability of which equals that of our first test, getting it right 999 times out of a thousand, but the errors of which are probabilistically independent of those of the first test. The second test will give a positive result in (almost) all the 300 true positive cases. It will also give a false positive result in 300 of the 300,000 false positives from the original test. So, we end up with 600 positives, of which half are real. The probability of having the disease, given positive results on both tests, is about .5. The second test, or “witness,” if you like, raises the probability from one in a thousand to an even bet.

The same principle applies to the case of Biblical miracles. Given our cautious estimates, it would take ten witnesses to make the miracle have close to .5 probability (actually, .4749), and twelve (!) to make it highly likely (.9888). Given our incautious estimates—appropriate for the most trusted disciples, such as Peter, James, and John, say—these levels are reached much more quickly. Five independent witnesses give the miracle an even chance of occurring; six make it highly probable.

One might object that the disciples are not independent witnesses, but very much under one another's influence; that the four Gospels are not entirely independent, but depend on many of the same sources; that many miracle reports were recorded long after the miracles are supposed to have taken place; and so on. There is something to these objections, though less, perhaps, than many think. Minor differences in the Gospel accounts offer evidence of independence. All the Apostles who faced imprisonment, beatings, and martyrdom for their testimony had strong incentive to recant anything for which they did not have overwhelming independent evidence. The Gospels appear to have been written within the lifespans of those who knew Jesus and witnessed the events recorded in them. But return to Paul's argument concerning the Resurrection. Writing perhaps just twenty to twenty-five years after the event, he points to hundreds of witnesses. Not all are independent, but many are. The credibility he attaches to the Resurrection is thus, reasonably, very high, even setting aside his own experience on the road to Damascus.

Paul was in a far better epistemic state with respect to Christ's Resurrection than we are, say, with respect to the attack on a canoeing President Carter by a crazed, swimming rabbit in 1980. That was surely an improbable event, a little further removed in time, witnessed by only a few government employees whose reliability may not compare very well with that of the disciples. Yet most of us—rationally—believe that it occurred. If we are to throw out belief in Biblical miracles on Humean grounds, we should throw out many of our historical beliefs on those very same grounds, for they would fail Hume's test too, and for the very same reasons.

Whether the New Testament provides us with adequate historical evidence of independent witnesses is, in any case, a matter of assessing the **(p.31)** actuality premise. I cannot enter into the historical discussion here. My point is simply that the argument from miracles could provide rational justification for accepting the occurrence of miracles if a relatively small number of independent witnesses testified to them.

So far, keep in mind, I have been discussing the toughest case: resurrection from the dead. The more typical miracles of healing involve events whose prior probability is far higher, since spontaneous remission, psychosomatic illness, and so on are well documented. Suppose such a probability is not one in ten billion but one in a million. Then, even on our cautious estimates, a miracle has an even chance if there are six independent witnesses, and is highly likely (.99) if there are eight. If the prior probability is one in a thousand, then, again on our cautious estimates, it takes only three independent witnesses to give an even chance, and five to make the probability .99.

Although most contemporary philosophers have overlooked the power of multiple witnesses in establishing the rationality of belief in miracles, an early critic of Hume, William Paley, noted it eloquently:

> If twelve men, whose probity and good sense I had long known, should seriously and circumstantially relate to me an account of a miracle wrought before their eyes, and in which it was impossible [I would prefer to say "unlikely"] that they should be deceived [i.e., they have high reliability]; if the governor of the country, hearing a rumor of this account, should call these men into his presence, and offer them a short proposal, either to confess to the imposture, or submit to be tied up to a gibbet [thus greatly reducing the probability of their testifying falsely, i.e., $l_K = P[T/(-M \& K))]$; if they should refuse with one voice to acknowledge that there existed any falsehood or imposture in the case; if this threat were communicated to them separately, yet with no different effect; if it was at last executed; if I myself saw them, one after another, consenting to be racked, burnt, or strangled, rather than give up the truth of their account; —still, if Mr. Hume's rule be my guide, I am not to believe them. Now I undertake to say that there exists not a skeptic in the world who would not believe them, or who would defend such incredulity.[25]

From a Bayesian point of view, Paley has it exactly right. If one person were to behave this way, we could always chalk it up to mental illness. Two? A remarkable coincidence, hard to understand. Three? Surely they must have seen something. Twelve? What kind of evidence do you want?[26]

(p.32) V. Transfer and Acquisition: Series

The second point I want to stress is that Biblical miracles rarely occur in isolation. They occur in patterns. The central miracles of the Judeo-Christian tradition occur in series. They are not isolated events. In fact, in the Scriptures miracles are relatively rare. There are three main series of miracles: (1) the miracles of the Exodus from Egypt, comprising the plagues, the escape across the Red Sea, the provision of water and manna, and the Ten Commandments; (2) the miracles of Elijah and Elisha; and (3) the miracles of Jesus and the Apostles.[27] George Schlesinger contends that the occurrence of miracles in series makes no difference: "additional reports do not raise the chance that at least one miracle has occurred."[28] But that is incorrect. In fact, they raise the chance that each such miracle has occurred.

Moses, Elijah, Elisha, Jesus, and the Apostles all work series of miracles. Sorensen and Holder argue that this counters Hume's argument, for we might have evidence that at least one miracle in a series occurred even if we cannot establish which one. I want to argue for a stronger point: each miracle in a series, and each miracle report attesting to it, lends evidential support to *all* the miracles of the series.

Return to the medical diagnosis analogy. You have been diagnosed, by a reliable test, as having a very rare illness. As we have seen, the likelihood that you have the disease is nevertheless quite low (in our example, one in a thousand). We might administer a second, independent test (or simply rerun the first, if errors are random); a positive result on that test would raise the probability of your having the disease to roughly .5, and a third, independent, positive report would make the probability .999.

But suppose that a second test is not available, and suppose that errors on the one test we have are not random, but caused by complex individual variations for which we cannot easily screen. Is there anything we might do to gain more information about the likelihood of having the disease? Of course. We might examine your medical history carefully for signs that would raise your prior probability of getting the disease. We might watch **(p.33)** you carefully for symptoms associated with the disease. In short, we might look to your history and to your future for anything correlated with the disease. If we find a family history of the disease or an earlier episode of a related illness, we increase our assessment of your likelihood of having the disease. We do the same if you begin to display characteristic symptoms.

Consider a different analogy. Suppose John reports that Mark hit a baseball 500 feet. Knowing that such an event is very unlikely, I refuse to believe it. And unfortunately there are no other witnesses I can ask. But I can see whether anyone has reported Mark doing something similar in the past, and wait to see whether there are reports of his doing it in the future. If I find other occasions on which Mark has hit a baseball very far, or performed other astounding feats of strength, I increase the credibility that I assign to John's report.

In the case of miracle reports, I submit, we can do exactly the same thing. We may or may not be able to find other witnesses; some miracles (such as the Transfiguration) have a few witnesses, others (the burning bush, for example) just one. Faced with a report, or collection of reports, we can look to the history of the person performing the miracle to see whether there are other reports of that person performing miracles. We can wait to see whether similar things occur in the future. If there are other reports, that may increase the probability that the report we start from is true.

Here is an advantage of my definition of miracles as events the best explanations for which involve supernatural agency. The occurrence of improbable events in series does not always increase the credibility of reports of each event. Suppose that a friend tells me that he won the lottery. His winning the very next week does not lend additional support to the report of his winning last week, for every play of the lottery is independent of every other. Because miracles lend rational support to belief in supernatural agency, however, they are not independent in that way. Series of unlikely events can be purposive, and can give us grounds for seeing the entire series as consisting of purposive acts of a single agent.

I am thus arguing for something stronger than Sorensen and Holder's point that, if there are several miracles supported by independent testimony, the probability of the disjunction of the miracle reports—that is, the claim that at least one of the miracles occurred—may be high even if the probability of each miracle, considered in itself, is very low. It is not just that, if we find reports of a series of miracles, there may be a significant probability that at least one happened. My point is that the miracles in the series may set a context for each other. The Egyptian who hears Moses threaten a plague of flies quite reasonably laughs; how can the acts of Moses or the Pharaoh cause that? But after seeing the flies, the boils, the illnesses, the hail, and the locusts, an Egyptian who does not tremble at the threat to the first-born is a fool. **(p.34)** Similarly, one might reasonably balk at believing that Jesus turned water into wine. But after hearing about the healings, the feeding of the five thousand, and the resurrections, that first miracle no longer seems so surprising.

In general, we assess prior probabilities and reliabilities in context. Most writers who give a contemporary reconstruction of Hume's argument suppress reference to context, assuming that it stays constant and makes no difference to the results. But that understates the credibility of miracles considerably. In assessing the credibility of reports of the Passover, for example, we must estimate the prior probability of the event. But, in so doing, we must ask, not “What is the likelihood of a plague striking down all the first-born overnight, leaving others, and those who have painted their doors with lamb's blood, unharmed, given that Moses predicts it?” but “What is the likelihood of such a thing, given that Moses predicts it—and that flies came as Moses predicted, boils came as Moses predicted, illnesses came as Moses predicted, hail came as Moses predicted, locusts came as Moses predicted, darkness came as Moses predicted, etc.?” Which way would you bet? The probability may not be high, but it is surely significant. The prior probability in context is much greater than the prior probability would be in isolation.

In assessing the credibility of reports of Christ's Resurrection, similarly, we must ask, not “What is the likelihood of someone's being raised from the dead?” but “What is the likelihood of someone who has been born of a virgin mother, displayed exceptional wisdom, healed the sick, fed the five thousand, walked on water, and raised others from the dead, being raised from the dead?” The probability is still not that high—consider the surprise, indeed, the shock and even fear of the women and the disciples on finding an empty tomb on Easter morning—but it is surely much higher than one in a billion.

Like multiple witnesses, then, the occurrence of miracles in series can have dramatic effects on the credibility of miracle reports. In this case, however, it does so by altering the context and thus affecting prior probability assignments. Recall that, if the prior probability of an event is one in ten billion, 10^{-10}, then, under our cautious estimates, ten independent witnesses are required to raise the probability to .5. Twelve are needed to raise it to a near certainty. If the prior probability is one in a million, then only six witnesses are needed. In general, each order of magnitude in prior probability corresponds to a witness. So, if the prior probability rises to one in a thousand, three witnesses suffice. If it rises to one in a hundred, two are enough. If it rises to one in ten, then even one witness report makes the event more likely than not to have occurred.

To see why this is so important, consider a miracle such as the Resurrection, to which we assign a prior probability of 10^{-10}. Suppose that occupies a place in a series of miracles, such as the turning of water into wine, the feeding of **(p.35)** the five thousand, and so on, in decreasing order of probability. Say that the first miracle has a prior probability of one in a hundred, and each further miracle in the series has the same prior probability given the context comprising the previous miracles of the series. Then the miracle could appear as early as fifth in the series. If, moreover, each miracle in the series has at least three witnesses, even the miracle with prior probability of 10^{-10} is overwhelmingly likely to have occurred. Each witness to a miracle earlier in the series raises the probability of the final miracle's occurrence.

I am not, of course, arguing for the absurd thesis that the probability of a conjunction is sometimes higher than the probability of each conjunct. I am arguing that, sometimes, the probability of a conjunction, given the evidence for all conjuncts taken together, is higher than the probability of some conjunct given a subset of that evidence. For any p and q, $Pr(p \wedge q) \leq Pr(p)$; but it may happen that $Pr((p \wedge q)/(e_p \wedge e_q)) \rangle Pr(p/e_p)$. Reports of miracles can provide evidence for other miracles directly or indirectly, by providing evidence of agency or power, for example, in a way that alters the context for other miracle reports. Far from being absurd, this is a special case of the insight that lends plausibility to coherentist accounts of justification.

One advantage of my account is that it explains a widespread ambivalence that many religious people feel about reports of miracles. We doubt whether Jonah was really swallowed by a large fish. We do not put much stock in claims to private mystical experience. We express skepticism about faith healing. We puzzle at people who see Mary's face on a tree. We laugh when, in a dream, God tells Homer Simpson, “Well, Homer, I'd better be going. I'm scheduled to appear on a tortilla in Mexico.”

Yet Christian faith rests, at least in part, on reports of miracles found in the Scriptures. Those miracles, moreover, fit together in such a way as to support the conclusion that they are purposive acts of a single supernatural agent. From the perspective of the historical theory, the miracles that deliver Israel from the bonds of Egypt form an essential part of the relationship between God and His chosen people. The miracles that Elijah and Elisha perform continue that relationship, keeping faith alive under the dissipating pressures of the cults of Ba'al and the crises of the house of Ahab. Most crucially, the miracles of Jesus's birth, ministry, and Resurrection, and the subsequent miracles of the Apostles, help define what it is for Jesus to be the Christ, the Son of the living God. Without those miracles, what is left of Christian faith? We may not go so far as Paul, who says that “If the dead are never raised to life, let us eat and drink, for tomorrow we die” (I Corinthians 15:32). But if we have no evidence for miracles, we have no evidence of God's interaction with Israel, with the world, or with us. We have no basis for any *de re* attitudes toward God at all.

(p.36) A typical Christian, then, tends to reject miracle reports in isolation but accept the series of miracle reports that constitute the backbone of the faith, seeing them, quite reasonably, as purposive acts of a single supernatural agent. Incoherence? Hypocrisy? Not at all. It is an entirely rational thing for a Bayesian Christian to do.

VI. Acquisition

So far I have followed the literature in concentrating on the issue of transfer rather than acquisition. I have been arguing, contra Hume, that it can be rational to invoke supernatural agency in response to miracle reports when they are attested by multiple witnesses or occur in series. But what about acquisition? Is it ever rational to think that an experience one is undergoing or has undergone is miraculous? Are there really events the best explanation for which might invoke supernatural agency? Specifying kinds of events that would have this property is difficult; to accomplish it would be to specify the range of possible miracles. We can at least say, however, that, normally, miracles are events that would not have occurred in the natural order without supernatural intervention, and that exhibit evidence of purposiveness. The miracles in the Biblical series I have discussed all fit that characterization. They are acts the best explanations for which involve supernatural agency, in fact, the agency of a single supernatural agent.

There are two arguments against a principle of acquisition, contending that it could never be rational for even the person undergoing it to describe an experience as a miracle. I shall argue that both arguments fail.

The first argument follows Hume in arguing that we could never acquire a rational belief in supernatural agency on the basis of any available evidence. It would always be more rational to doubt the testimony of one's senses—to suspect illusion, hallucination, madness, confusion, dreaming, or some similar epistemic dysfunction—than to accept supernatural agency or the evidence that points to it. This is Hume's argument applied to the senses rather than to witnesses. It has the merits and the demerits of that argument. Supposing that the senses are normally but not perfectly reliable, and that miracles are extremely unlikely, the probability that what one seems to be experiencing is really happening may be very low.

As I have argued, however, that argument is not decisive. We can respond just as we do when we doubt our senses for other reasons. Sometimes we can check one sense with another—reach out to see whether our sense of touch confirms what we see, for example. We can check our senses against **(p.37)** those of others if there are other witnesses to the same events. “Do you see what I see?” “Did you hear that?” We can moreover see whether the extremely unlikely event is singular or whether it is followed by other unlikely events. Just as the testimony of witnesses can constitute strong evidence in favor of a miracle report if there are multiple witnesses or multiple miracles, so the evidence of the senses for a miracle can be quite strong if corroborated by others or by our own senses on other occasions.

Again, the Humean argument seems to go too far. It argues against belief in extremely unlikely events solely because they are extremely unlikely. So it seems to indict belief in the occurrence of any extremely unlikely event.

Suppose that you buy a lottery ticket. The probability of your winning the lottery is extremely low. But it seems absurd to suggest that nothing you could experience could provide rational justification for thinking that you won the lottery. Given any one bit of evidence, let's say, the probability that you are suffering an illusion, a hallucination, a dream, madness, etcetera, is significantly higher than the probability that you have actually won. But it would be preposterous to conclude that you could never be in an epistemic position to conclude that you had won the lottery. You could check the ticket again. You could ask someone else to check the numbers. You could hand the ticket over to the officials in charge to see how they behave. You could see the bank statement after the money was deposited into your account. No one thing could be enough to support rational belief that you had won. But the accumulated evidence could.

A Humean argument against acquisition thus fails for exactly the reason the argument against transfer fails. We can find support for a belief in supernatural agency in the same ways in which we can seek support for believing that other highly improbable events have occurred. The women who find Jesus's tomb empty are “bewildered” and run to get others (Matthew 18; John 20); Thomas, doubting the testimony of others, wants confirmation from his own senses: “Except I shall see in his hands the print of the nails, and put my finger into the print of the nails, and thrust my hand into his side, I will not believe” (John 20:25). Subsequent appearances of Jesus seem, quite reasonably, to increase the confidence of the disciples. All this is entirely rational on Bayesian grounds.

The second argument would depart from Hume, arguing against belief in miracles not because their occurrence is extremely unlikely but because they involve attributions of supernatural agency. One might argue that nothing could count as evidence of a supernatural purpose—that nothing could count as a sign of God's activity in the world. Now, it is hard to see how to construct such an argument in general, without relying either on the metaphysical impossibility of miracles (as putative violations of universal and **(p.38)** necessary laws, say) or on the first argument that focuses on rational belief. But the argument makes a sweeping claim, placing all the following examples in the same position from an epistemic point of view. Admittedly, purported signs of God's purpose are often ambiguous. But there appear to be significant evidentiary differences between the following fictional cases, drawn from *The Simpsons*:

- Homer, in *Lisa's Sax*: “Musical instrument? Could that be a way to encourage a gifted child? [to the heavens] Just give me a sign! [At that moment, the store owner happens to put a sign in the window reading “Musical Instruments: The Way To Encourage A Gifted Child.”] Eh, it works for me.”
- Homer, in *Homer Bad Man*: “You mean…I'm on my own? I've never been on my own! Oh no…on own…on own! I need help…oh, God, help me. Help me, God!” [The phone rings; Homer answers it very slowly.] Homer: [very slowly] “Y'ello?” Man: “Hello, Homer. This is God…frey Jones from the TV magazine show ‘Rock Bottom’.”
- Lisa, in *Homer the Heretic*, surveying a raging fire at the Simpson house: “Truly this was an Act of God.” The fire spreads to Ned's House. Homer: “Hey. Flanders is a regular Charlie Church, and God didn't save his house.” A tiny cloud forms over the Flanders House; the rain douses the fire, and the damage is sealed with a rainbow. Homer: “D'oh!”

> - Ned, in *A Star is Burns*: “Now, Maude, in our movie you lay Moses in the basket, then put it among the reeds, OK? Lights, camera, ac-diddelydoddely-doodely-action, Jackson!” [Maude puts the basket in the water; it is quickly swept away.] Todd: “Help meeee…eeee… eeee…” [The sound vanishes as Todd passes behind some trees.] Ned: “Flanders to God, Flanders to God, get off your cloud and save my Todd!” [Lightning fells a tree across the river, blocking Todd's path.] Everyone: “Yay!” Ned: “Thanks, God!” God: [making the OK sign through the clouds] “Okily dokily!”

If the first two examples seem relatively easy to explain naturalistically, the last two do not. Moreover, they seem to exhibit obvious signs of purposiveness. They appear to satisfy ordinary criteria for attributing agency. The burden is surely on the opponent to show why no such case could possibly count as evidence of supernatural agency.

I do not have a criterion by which to distinguish cases like the first two from those like the last two. The last two examples are of course far less likely, within the natural order, than the first two. They offer stronger **(p.39)** evidence of purposiveness. They are harder to interpret as mere coincidences. These are relevant factors, but far less than a criterion.

We should not expect, however, to be able to draw a sharp line between cases that would provide evidence for supernatural agency and those that would not. All that we need is the possibility of cases that would. Bertrand Russell was once asked what he would say if, after death, he found himself confronted by God, who demanded to know why he did not believe in Him. “Not enough evidence, God! Not enough evidence!” was his reply. Note that he did not say, “There *could not have been* any evidence.” Nor did he say, “What evidence do I have that I am talking to God *even now*?” Those answers sound specious. But they are exactly the answers the argument under consideration would demand.

There is a strong analogy, moreover, between these kinds of cases and the question of whether a physical system is closed. Within a physical system, it is possible to have evidence of various kinds (e.g., of the effects of acceleration throughout the system) that external forces are operating—indicating, in other words, that the system is not closed. The same is true of ecological systems, to take another kind of case, in which there can be empirical evidence for or against the operation of external forces on the system. There is nothing intrinsically mysterious about arguments that given systems are or are not closed. There appears to be nothing mysterious, therefore, about arguments that the order of nature itself is or is not closed. There can be evidence in favor of supernatural agency just as there can be evidence of interference, even intentional interference, with physical or ecological systems.

VII. Conclusion

I have argued that, although Hume shows it would be irrational to accept a single report of a single, isolated occurrence of a violation of a law of nature, nothing at all follows about the credibility of the central Biblical miracle stories. Multiple witnesses attest those stories, and they fall into series, supporting one another. Hume's argument shows that believing in isolated violations of laws of nature on the basis of a single witness's testimony flies in the face of reason. Skepticism about claims of isolated miracles, mystical experience, or privately transmitted visions is justified. But what is right in Hume's argument is fully compatible with rationally believing that God revealed Himself to Moses, sent him to Pharaoh, sent the plagues on Egypt, brought the Israelites out of Egypt, rescued them from Pharaoh's army, and led them (circuitously!) to the promised land. It is fully compatible with rationally believing that Elijah trounced the priests of Ba'al at Mount **(p.40)** Carmel or that Elisha raised the son of the Shunammite woman from the dead. And it is fully compatible with rationally believing that Christ was conceived by the Holy Ghost, was born of the Virgin Mary, and rose from the dead on the very first Easter morning.

References

Bibliography references:

Dawid, P. and Gillies, D. (1989), "A Bayesian Analysis of Hume's Argument Concerning Miracles," *Philosophical Quarterly* 39: 57–65.

Earman, J. (1993), "Bayes, Hume, and Miracles," *Faith and Philosophy* 10: 293–309.

Holder, R. D. (1998), "Hume on Miracles: Bayesian Interpretation, Multiple Testimony, and the Existence of God," *British Journal for the Philosophy of Science* 49: 49–65.

Hume, D. (1748), *An Enquiry Concerning Human Understanding*. All quotations in the text are from Chapter X, "Of Miracles", identified by section number.

McGrew, T. and McGrew, L. (2009), "The Argument from Miracles: A Cumulative Case for the Resurrection of Jesus of Nazareth," in W. L. Craig and J. P. Moreland (eds.), *Blackwell Companion to Natural Theology* (Oxford: Blackwell).

Millican, P. (1993), "Hume's Theorem Concerning Miracles," *Philosophical Quarterly* 43: 489–95.

Nozick, R. (1974), *Anarchy, State and Utopia* (New York: Basic Books).

Owen, D. (1987), "Hume versus Price on Miracles and Prior Probabilities: Testimony and the Bayesian Calculation," *Philosophical Quarterly* 37: 187–202.

Paley, W. (1794), *Evidences of Christianity* (Philadelphia, PA: J. J. Woodward), reprinted 1836.

Popkin, R. H. (2003), *The History of Scepticism from Savonarola to Bayle* (Oxford: Oxford University Press).

Randall, J. H. (1926), *The Making of the Modern Mind: A Survey of the Intellectual Background of the Present Age* (Boston: Houghton Mifflin).

Schlesinger, G. (1987), “Miracles and Probabilities,” *Noûs* 21: 219–32.

Sobel, J. H. (1987), “On the Evidence of Testimony for Miracles: A Bayesian Interpretation of David Hume's Analysis,” *Philosophical Quarterly* 37: 166–86.

—— (1991), “Hume's Theorem on Testimony Sufficient to Establish a Miracle,” *Philosophical Quarterly* 41: 229–37.

Sorensen, R. A. (1983), “Hume's Scepticism Concerning Reports of Miracles,” *Analysis* 43: 60.

Stephen, Leslie (1903), *An Agnostic's Apology, and Other Essays* (London: Smith, Elder & Co.).

Swinburne, R. (1979), *The Existence of God* (Oxford: Clarendon Press).

—— (2003), *The Resurrection of God Incarnate* (Oxford: Oxford University Press).

Notes:

(*) I delivered earlier versions of this paper as the George S. Heyer, Jr., Lecture at Austin Presbyterian Theological Seminary and at a Philosophy of Religion conference at Baylor University. I am grateful to my audiences and to two anonymous referees for their questions, criticisms, and advice. I owe a special debt to Robert C. Koons for encouragement and helpful discussions.

---

(1) I say “generally classical” because the arguments vary in what they require. The cosmological argument understands God as the creator, first cause, or necessary being. The ontological argument understands God as the perfect being, the greater than which cannot be thought. The teleological argument understands God as the designer of the universe. As Kant stresses, only the ontological argument appears to offer much hope for tying together the properties classically attributed to God. There is no obvious argument from being the creator to being the designer, or vice versa; being an architect is independent of being a builder. Still less does either property entail perfection and its implications (omnipotence, omniscience, or moral goodness, for example). Kant may be too optimistic about the ontological argument's ability to do this; does perfection entail creation? In any case, whether the argument from miracles is able to tie together the properties of God is something I shall discuss later in this paper.

(2) For excellent summaries of some of the issues, see Swinburne (2003) and McGrew (2009). There is a larger point lurking here. Many contemporary intellectuals think religion runs afoul of sophisticated, modern, scientific thinking. The rationale for that view depends largely on nineteenth-century and early twentieth-century developments—most notably, higher criticism, sociological and anthropological theories of religion, and Freudian psychology—that on careful examination today appear anything but scientific. Far from being sophisticated, in short, this sort of intellectual hostility to religion may be a remnant of outmoded conceptions of what counts as scientific. But defending that claim goes far beyond the scope of this paper.

(3) Randall (1926: 293).

(4) Stephen (1903) and Sobel (1987).

(5) Sorensen (1983).

(6) From now on, for simplicity, I shall speak solely of events.

(7) See Nozick (1974: ch. 7). The analogy suggests the possibility of a third condition, of rectification or, in this context, reformation of rational credence of supernatural agency. I think there does indeed need to be such a condition, since chains of belief may be corrupted in a variety of ways, but spelling it out goes beyond the scope of this paper. Formulating a reformation condition would nevertheless be an extremely important thing to do. Indeed, one can read many of the figures on both sides of Reformation-era debates as trying to formulate such a condition or argue that such a task is impossible on pain of skepticism. See Popkin (2003).

(8) A reformation condition, of course, would lead us to add a third condition—that one's epistemic situation might include grounds for rational belief on the basis of reformation of earlier beliefs.

(9) The series would, of course, not have to consist simply of burning bushes; the kinds *A* and *B* might be described more abstractly, and, in Moses's case, would include turning the rod into a snake and back again, predicting the plagues, and so on.

(10) Hume (1748: 90).

(11) Giving a general account of agency, much less supernatural agency, goes beyond the scope of this paper. I take it, however, that, on any reasonable account of agency, being able to be seen as purposive will count in favor of ascribing an event to agency.

(12) There is a subtle issue here about the meaning of counterfactuals and the status of unlikely events. Would an extremely unlikely event count as something that would not happen in the course of nature? It would not normally happen, of course. A similar point could be made about quantum tunneling. We might not know whether the event would have occurred in the order of nature; we might even wonder whether that question has any determinate answer.

(13) I see whatever emotional power these arguments have as resting on the intuition that the universe's existence is miraculous. One could accommodate this intuition in the McGrews' definition by saying that the upshot of the cosmological, anthropic, and teleological arguments is that the natural order itself would not exist in the natural order of things, for the natural order depends in one way or another on the supernatural.

(14) Hume (1748: 86).

(15) This argument strikes me as unpersuasive for another reason. If we take laws of nature as strictly universal and necessary, the defender of miracles can simply deny that there are any true laws of nature.

(16) Hume (1748: 87).

(17) Ibid. (89).

(18) Ibid. (98).

(19) Swinburne (1979, 2003), Sobel (1987, 1991), Owen (1987), Schlesinger (1987), Millican (1993), Dawid and Gillies (1989), among others.

(20) All calculations below use Bayes's formula, not this approximation, though the approximation makes deriving answers very close to the actual values extremely easy.

(21) This figure is somewhat arbitrary, but its actual value, within a very broad range, makes relatively little difference. McGrew (2009) estimates the Bayes factor, given all available evidence concerning the resurrection, at 10^{43}; if that is correct, then only values less than 10^{-43} would have a significant impact on the argument to follow.

(22) Perhaps surprisingly, this value makes little difference to the calculation when p_K is small, for it occurs in both numerator and denominator multiplied by p_K.

(23) Earman (1993) and Holder (1998).

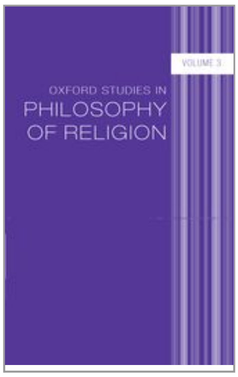
(24) Cf. Millican (1993).

(25) Paley (1794: 6).

(26) An anonymous reviewer objects that Paley presents evidence of strong conviction, but not of independence. The Apostles met their fates, however, separately and alone, which seems to me evidence of independence in a way that collective torture and execution might not be.

(27) I am setting aside two important kinds of divine interaction that run throughout Scripture but deserve their own treatment: divine inspiration and prophecy. If the canon is correct, every book of the Bible is an instance of divine inspiration. And prophecy is important not only to the books of the Prophets but also to much of the remainder of the Bible. Both inspiration and prophecy arguably involve miracles. But they are special in ways I cannot treat adequately here. I will simply note that repeated instances in these cases seems to make no epistemic difference, perhaps because there are easily available naturalistic explanations for people writing books and making predictions that turn out to be true.

(28) Schlesinger (1987: 230).



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

Omniprescience and Tough Choices

E. J. Coffman

DOI:10.1093/acprof:oso/9780199603213.003.0003

Abstract and Keywords

Several philosophers have argued that an omniprescient person could not act in certain ways. This chapter presents and assesses what may be the most promising such argument: the Tough Choices Argument (TCA). The TCA concludes that an omniprescient person could not make a tough choice—i.e., a choice between alternative courses of action in light of knowledge that none is uniquely reasonable. The TCA derives from an argument due to Tomis Kapitan; and the objection to the TCA developed in this chapter superficially resembles a reply to Kapitan's argument due to David Hunt. Along the way, then, the chapter discusses Kapitan's argument, and Hunt's reply, to show how they differ (respectively) from the TCA and the objection to it developed herein.

Keywords: omniprescience, omniscience, God, deliberation, choice, decision, intentional action, freedom, theism, determinism

Let's say that you are *omniprescient* iff you always believe—occurrently and with maximal confidence—all and only truths, *including* ones about the future. Several philosophers have argued that an omniprescient person couldn't act in certain ways.[1] In this paper, I'll present and assess the most promising such argument I know of: the *Tough Choices Argument (TCA)*. The TCA concludes that an omniprescient person couldn't make a *tough choice*—i.e., a choice between alternative (mutually excluding) courses of action in light of knowledge that none is uniquely reasonable.[2] The TCA derives from an argument due to Tomis Kapitan;[3] and my favored objection to the TCA superficially resembles a reply to Kapitan's argument due to David Hunt.[4] Along the way, then, I'll discuss Kapitan's argument, and Hunt's reply, to show how they differ from the TCA and my favored objection to it.

To motivate my presentation and assessment of the TCA, I'll start by arguing that many traditional theists are committed to thinking that God is an omniprescient person who makes tough choices—and so, are committed to denying the TCA's conclusion. Most traditional theists think God is free **(p.42)** with respect to which feasible world is the actual world.[5] Now, if there's a single best feasible world, then God *isn't* free with respect to which feasible world is the actual world (more on this—call it **Questionable Step 1 [QS1]**—below). So, most traditional theists should deny that there's a single best creatable world. But if there's *not* a single best creatable world, then God makes some tough choices (more on this—call it **Questionable Step 2 [QS2]**—below). So, most traditional theists should think that God makes some tough choices. Further, most theists who should think God makes some tough choices will also think God is omniprescient. But the TCA threatens to rule out the possibility of an omniprescient person who makes tough choices. The TCA thus poses a serious threat to traditional theism. Accordingly, it deserves the attention of anyone interested in the viability of such theism.

I flagged two questionable steps in the above argument: QS1 and QS2. There are strong arguments available for each of these claims. Let's start with QS1. Suppose there's a single best feasible world, BW. BW's status as best feasible world is presumably determined by which worlds are feasible and their associated overall values (which perhaps supervene on other facts about the worlds). Plausibly, God isn't free with respect to which worlds are feasible *or* the associated overall values of those worlds. So BW's status as best feasible world is entailed by facts relative to which God isn't free. But one lacks freedom relative to logical consequences of facts relative to which one lacks freedom.[6] So, God isn't free with respect to the fact that BW is the best creatable world. Finally, it's plausible that God must actualize the best creatable world provided there is such a thing.[7] So, God's actualizing BW is entailed by facts relative to which he lacks freedom. But then, by the above closure principle, God isn't free with respect to which feasible world becomes the actual world. By conditional proof, we now have QS1.[8]

Now for QS2. Suppose there's *not* a single best creatable world. Then *either* there's an infinite hierarchy of better and better feasible worlds *or* there's a tie for best feasible world *or* there are incommensurable feasible worlds.[9] Suppose one of these options holds. Then each creatable world is presumably such that **(p.43)** God knows his reasons do not conclusively favor actualizing *it*.[10] So, when God chooses which feasible world to actualize, he knows that none of the considered options is uniquely reasonable. When God chooses which creatable world to actualize, he makes (what we're calling) a tough choice: God chooses between alternative courses of action in light of knowledge that none is uniquely reasonable. By conditional proof, we now have QS2.

I think the last couple of paragraphs suffice to justify QS1 and QS2, the two questionable steps of my argument that the TCA poses a serious threat to traditional theism. I hope that that argument convinces you (if you weren't convinced already) that the TCA deserves serious attention, thus motivating my presentation and assessment of it.

Without further ado, here is the Tough Choices Argument:

1. Suppose S is omniprescient (Assumption for Conditional Proof).
2. *If* a person makes a tough choice between alternative courses of action A and B, *then* (prior to the choice) the person believes about each of A and B that her taking it is epistemically possible—i.e., she believes about each of A and B that she *may* take it (where “may” here expresses an epistemic, as opposed to a metaphysical, modality).[11]
(p.44)
3. *If* S believes about each of A and B that she may take it, *then* S lacks a maximally firm occurrent belief as to whether she will take A. [1]
4. So: *If* S makes a tough choice between A and B, *then* (prior to the choice) S lacks a maximally firm occurrent belief as to whether she will take A. [2,3]
5. But: S *has* a maximally firm occurrent belief as to whether she will take A. [1]
6. So: S does not make a tough choice between A and B. [4,5]

C. So: *If* S is omniprescient, *then* S does not make a tough choice between A and B. [1–6]

(C) generalizes to the thesis that an omniprescient person can't make a tough choice. So, if the TCA is sound, being omniprescient excludes making tough choices.

Given the definition of “omniprescience” in play here, (1) clearly implies (5); and my later critical assessment of the TCA will focus on (2) and some of the main considerations in its favor. One good way to clarify the argument at this point is to show how (3) follows from (1). After doing that, I'll briefly discuss an argument from Tomis Kapitan that inspires the TCA, and an important reply to Kapitan's argument from David Hunt. This discussion will make clear how the TCA, and my favored objection to it, differ from Kapitan's argument and Hunt's reply. We'll then be ready to consider my objection to the TCA in detail.

How, exactly, is (3) supposed to follow from (1)? Well, assume (1) is true, and that S has a maximally firm occurrent belief as to whether she'll take course of action A (as opposed to B). *Either* S believes that she'll take A, *or* S believes that she'll take B. Suppose S believes she'll take A. Then S has a maximally firm occurrent belief that she'll take A, *and* (by 1) simultaneously lacks the belief that she'll take B. If S were now to (assertively) utter “I may take course of action B,” her assertion would be false: under these conditions, it's not epistemically possible for S that she'll take B. So, given that S has a maximally firm occurrent belief that she'll take A *and* simultaneously lacks the belief that she'll take B, it follows (by 1) that S doesn't regard her pursuing B as epistemically possible. Parallel reasoning shows that if S instead believes that she'll take course of action B, then S doesn't regard her pursuing A as epistemically possible. So, if S has a maximally firm occurrent belief as to whether she will pursue course of action A (as opposed to B), then S doesn't believe about each of A and B that she may pursue it. Contrapose that last claim and you have (3). So (1) does imply (3) after all.

I'll now explain how the TCA differs from the argument due to Kapitan that inspires it, and how an important reply to Kapitan's argument due to Hunt **(p. 45)** differs from my favored objection to the TCA. The main difference between the TCA and Kapitan's similar argument is that the latter (unlike the former) concerns not *making a tough choice* but instead simply *gaining an intention*. More fully, Kapitan employs the following premise that is somewhat similar to—but, as we'll soon see, much less plausible than—premise (2) of the TCA:

> *Kapitan's Premise (KP)*: *If* a person acquires an intention to (take course of action) A, *then* (prior to gaining the intention) the person believes both that she may A *and* that she may not A—i.e., prior to gaining the intention, she believes about each of A and not-A that it is epistemically possible.[12]

There are several counter-examples to KP that do not also impugn premise (2) of the TCA. To begin to see this, consider the following three cases of intention acquisition described by Alfred Mele:[13]

> 1. *The Front Door*: As a part of my normal routine, I unlock my front door. Further, I unlock the door intentionally, gaining an intention to now unlock the door just before unlocking it.
> 2. *The Beer Desire*: Walking home from work, Helen notices her favorite brand of beer on display in a store window. The sight of the beer prompts a desire to buy some, and her acquiring that desire issues straightaway in an intention to buy some.
> 3. *The Experienced Driver*: [Sometimes]…a perceptual event, given the agent's psychological profile, straightaway prompts an intention to A. Seeing a dog dart into the path of his car, an experienced driver who is attending to traffic conditions may immediately acquire an intention to swerve.

Mele contends that in these cases:

> there is no uncertainty that intention acquisition resolves. I was not uncertain about whether to unlock my door, Helen was not uncertain about whether to buy the beer, and the driver was not uncertain about what course of action to take. At no time were any of us uncertain about the matters at issue.[14]

On the basis of such examples, Mele concludes that it's possible for you to gain an intention to A without prior uncertainty as to whether you'll A.

Such examples justify a stronger conclusion. We could fill in the above cases so that—prior to gaining the relevant intentions—Mele was (psychologically) certain he'd unlock his door, Helen was certain she'd buy the beer, **(p.46)** and the driver was certain he'd swerve to miss the dog. Further, we can add that the subjects did not also believe that they wouldn't perform those acts. So amplified, the cases cast serious doubt on KP: each involves a person who gained an intention to A *yet* did not previously regard failure to A as epistemically possible. But since the people in such cases aren't making tough choices (i.e., choosing between incompatible courses of action in light of knowledge that none is uniquely reasonable), the cases do not also impugn premise (2) of the TCA.

For a final kind of counter-example to KP, consider the following case due to Dana Nelkin [my emphasis]):

> *Distraught Friend*: [Y]ou are engaged in a long-anticipated activity (e.g., watching an overtime period of a championship basketball game, attending a concert, taking a once-in-a-lifetime trek). You receive a call from a friend who desperately needs to talk to someone about the sudden and unexpected death of a family member. In the past, you have always [decided] in favor of talking to your friend; indeed, this is the kind of person you are. Based on these considerations, and perhaps others, you *know* you will decide the same way today. However, you haven't…[gained an intention] to do so yet. But you can and do.[15]

We can understand Distraught Friend so that your knowledge that you'll talk to your friend (which you have *before* gaining an intention to do) derives in part from your knowledge that your reasons conclusively favor talking to your friend. So understood, Distraught Friend is a counter-example to KP that does not also threaten premise (2) of the TCA. Provided you know that your reasons conclusively favor talking to your friend, then (by definition) your choice to talk to your friend is not a “tough choice.” And we can safely suppose that, prior to your choosing to talk to your friend, you did not regard failure to talk to your friend as epistemically possible. (If you had said—while knowing, perhaps with maximal confidence, that you were about to choose to talk to your friend—“I may just skip out on my friend this time,” you'd have said something false.)

(The above counter-examples to KP imply that your knowing, or being maximally confident, that you'll perform action A is compatible with your lacking an intention to A. Some philosophers have claimed that these are incompatible. Writes Carl Ginet: “[I]t is conceptually impossible for a person to know what a decision of his is going to be before he makes it.”[16] [cf. Hampshire and Hart (1958).] The standard and, I believe, decisive reply to this claim is that it confuses propositional attitudes that **(p.47)** are clearly logically independent: *knowledge*, or *psychological certainty*, and *intention*.)

Let's sum up the last few paragraphs. The main difference between the TCA and Kapitan's similar argument is this: the latter employs KP, whereas the former employs an analogous claim—premise (2)—that concerns not *simply gaining an intention* but instead *making a tough choice*. There are numerous counter-examples to KP that do not also threaten (2). So, what's arguably the TCA's weakest link is considerably stronger than what's arguably the weakest link of Kapitan's argument. This is a clear way in which the TCA improves upon the argument that inspires it.

My favored reply to the Tough Choices Argument will be that it fails because it depends on the alleged incompatibility of omniprescience and freedom. More on precisely *how* the TCA depends on the indicated incompatibility thesis, and *why* such dependence spells doom for the TCA, in a moment. What I need to do now is to distinguish my favored objection to the TCA from an important reply to Kapitan's similar argument due to David Hunt.

In Hunt's terminology, arguments like the TCA and Kapitan's argument are formulations of the “Problem of Agency”; and arguments for the conclusion that an omniprescient being would rule out free agency are formulations of what Hunt calls the “Problem of Freedom.” He argues that the Problems of Agency and Freedom “are really one, and the resolution of this joint problem depends on current investigations of the Problem of Freedom rather than the Problem of Agency.” As will become clear, I agree with Hunt's conclusion that the “Problem of Agency is parasitic on the Problem of Freedom.” But I reject his argument for our shared conclusion:

> The Problem of Agency presupposes the presumptive agent's acceptance of a fatalistic argument; but only the argument for Theological Fatalism stands any chance of being sound; therefore the Problem of Agency is ineliminable for a rational agent only if the argument for Theological Fatalism is sound. Since the latter argument sets forth the Problem of Freedom, the Problem of Agency (in its only interesting form) is dependent on the Problem of Freedom.[17]

The main reason Hunt offers for believing our shared conclusion is this: any argument that counts as an interesting version of the Problem of Agency will ascribe to the omniprescient agent in question a belief in some or other argument for the thesis that an omniprescient being rules out the existence of free agents.

(p.48) So far as I can see, though, neither the TCA itself nor the argument that results from replacing (2) with KP (Kapitan's argument) ascribes to the omniprescient agent in question (S) belief in some or other argument for the incompatibility of omniprescience and freedom. Indeed, we seem to have formulated and discussed both the TCA and Kapitan's argument without also laying out an argument that an omniprescient being rules out the existence of any free agents. Further, both the TCA and Kapitan's argument count as interesting versions of (what Hunt calls) the Problem of Agency. Thus, Hunt hasn't provided good reason to accept our common conclusion: we can formulate interesting versions of the Problem of Agency without also presenting and endorsing a version of the Problem of Freedom.

We're now ready to consider my favored objection to the TCA, which focuses on:

> 2. *If* a person makes a tough choice between alternative courses of action A and B, *then* (prior to the choice) the person believes about each of A and B that her taking it is epistemically possible.

After presenting my objection to the TCA, I'll justify its two questionable steps.

The best case for (2) doesn't favor (2) over the following somewhat different requirement on making a tough choice:

> *Belief in Freedom Requirement (BFR)*: *If* a person makes a tough choice between alternative courses of action A and B, *then* the person believes about each of A and B that she's (metaphysically) free to take it (alternatively, that she has it within her power to pursue it).[18]

In light of the (initially questionable—so see below) point that the strongest case for (2) is at best neutral between (2) and BFR, the TCA's proponent must replace (2) with the weaker:

> 2*. *If* a person makes a tough choice between alternative courses of action A and B, *then either* the person believes about each of A and B that her taking it is epistemically possible *or* she believes about each of A and B that she's (metaphysically) free to take it.

(p.49) Once (2) has been replaced with (2*), the TCA's proponent must (on pain of offering an invalid argument) replace (3) with:

> 3*. *If* (omniprescient) S *either* believes about each of A and B that her taking it is epistemically possible *or* believes about each of A and B that she's free to take it, *then* S lacks a maximally firm occurrent belief as to whether she'll take A.

Now, 3* has the following:

> *Obvious Consequence (OC)*: *If* (omniprescient) S believes about each of A and B that she's free to take it, *then* S *neither* believes that she'll take A *nor* believes that she'll take B.

And, finally, OC entails that the existence of an omniprescient being rules out the existence of free agents (a second questionable step—see below).

In sum: the TCA is tenable only if it employs a premise—viz. (3*)—that entails the incompatibility of omniprescience and freedom. So the TCA depends dialectically on the incompatibility of omniprescience and freedom. But then the TCA fails for at least one—for what it's worth, I'd say *all*—of the following three reasons.

First, since omniprescience is *compatible* with freedom, the TCA depends dialectically on a false proposition. Second, since (as the relevant debate presently stands) it's at least as likely as not that omniprescience is compatible with freedom, the TCA depends dialectically on a highly questionable thesis.[19] Third, since (in any case) the TCA clearly does depend dialectically on the incompatibility of omniprescience and freedom, the argument falls short of a key part of its billing—viz., that it's a *special, independent* problem for (the relevant kind of) theistic belief, a serious challenge to such belief regardless of whether omniprescience excludes free agency.[20]

What remains is to justify the two questionable steps in my objection. We'll take them in reverse order. First, why think OC entails that omniprescience is incompatible with the existence of free agents? Well, suppose that the existence of an omniprescient person (S) is compatible with the existence of free agents. Presumably, if freedom and omniprescience are compatible, then it's also possible that there be an omniprescient person who is also free. Notice that the TCA provides no reason at all to deny this plausible conditional *unless* the TCA is supplemented with the claim that acting freely **(p.50)** requires making tough choices. But that last claim is refuted by cases involving a free act whose agent knows her reasons conclusively favor some or other of her considered acts—a suitably modified version of Distraught Friend (see above) would fit this bill, as would certain examples involving "weakness of will."[21] And if it's possible that there be a free omniprescient person, then OC is false: it's possible that omniprescient S correctly believes about each of A and B that she's free to take it, yet *also* correctly believes about one of A and B that she will take it. So, by conditional proof, if omniprescience is compatible with freedom, then OC is false. Contrapose that last claim and we have the second of the two (initially) questionable steps of my objection to the TCA: OC entails that omniprescience excludes freedom.

Now for that first questionable step. Why think the strongest case for premise (2) of the TCA is at best neutral between (2) and BFR? In my judgment, the following three data jointly constitute the best available prima facie case for (2):[22]

1. *Datum 1 (D1)*: *If* you make a tough choice between alternative courses of action A and B, *then* (prior to the choice) you believe that it is (in some sense) possible for you to take either of A and B.[23]
2. *Datum 2 (D2)*: If a person who's trying to make a tough choice is asked whether he's aware of factors that threaten the freedom of the upcoming choice, the following would be a natural reply: "No, I'm not aware of any such factors. *So far as I can tell*, the outcome is up to me."[24]

3. *Datum 3 (D3)*: A Determinist (i.e., someone who believes in Determinism—the remote past and physical laws together fix a unique future) could make a tough choice without thereby holding explicitly inconsistent beliefs.[25]

(p.51) Because my objection to the TCA doesn't require the claim that BFR explains D1–D3 *better* than (2) does, I'll just grant that (2) provides a good explanation of D1–D3. What I need to argue is that BFR can explain D1–D3 *at least as well as* (2) does. Clearly, each requirement straightforwardly implies D1. So our question becomes this: can BFR provide a good explanation of D2 and D3?

Admittedly, there's prima facie tension between D2 and BFR. On BFR, someone who makes a tough choice believes (prior to the choice) that she is free (has it within her power) to take any of her considered courses of action. Initially, then, it looks as though BFR's advocate should predict that someone who's trying to make a tough choice will *unqualifiedly* self-ascribe freedom relative to the upcoming choice when asked whether she's aware of any “freedom-threatening” factors. But that initial appearance is misleading: BFR's proponent can make perfect sense of such a person's passing up an unqualified self-ascription of freedom for something weaker in circumstances of the relevant sort. Here are two good “BFR-friendly” explanations of D2.

On the one hand, the question whether you're aware of “freedom-threatening” factors may distract you away from the deliberative task at hand, resulting in a temporary suspension of your deliberation. Provided this happens, you're no longer in the process of trying to make a tough choice when fielding the indicated question. But if you've stopped deliberating to answer the question, then—consistent with BFR—you may no longer believe you're free to do multiple different things. And your no longer holding that belief would explain why you pass up an unqualified self-ascription of freedom for a qualified one (e.g., “*So far as I know*, I'm free with respect to the outcome…”).

On the other hand, you might believe throughout the entire episode that you are free to do each of multiple different acts. However, you might naturally take the question “Are you aware of any ‘freedom-threatening’ factors?” as a bit of evidence that such factors are present, and so as a bit of evidence that you're *not* free to do multiple different things. Under such circumstances, you might naturally hesitate to make an unqualified self-ascription of freedom: you might naturally feel that such an assertion would be conversationally inappropriate in such circumstances—i.e., circumstances where you take yourself to have been offered a bit of evidence for thinking you *aren't* free with respect to the outcome.

Consider an analogy. You go to see the doctor, justifiedly confident that you're getting enough exercise. Upon seeing you, the doctor immediately asks: “Are you getting enough exercise?” It would be natural for you to regard the question itself as a bit of evidence that you're *not* getting enough **(p.52)** exercise, though not enough evidence to require you to immediately drop your belief that you *are* getting enough exercise. So, while you remain confident (for the moment anyway) that you *are* getting enough exercise, you take yourself to have been offered a bit of evidence that you're not. In such a situation, it would be natural (and perhaps even conversationally required) to bypass an unqualified “Yes” for something weaker—e.g., “*So far as I know*, I'm exercising enough…” or “I *think* I'm exercising enough…”. Similarly, it would be natural for a person who is asked, while trying to make a tough choice, whether she's aware of any “freedom-threatening” factors, to bypass an unqualified self-ascription of freedom for something weaker—*even if* she remains justifiedly confident throughout that she's free with respect to the outcome of her deliberation.

In light of the last couple of paragraphs, I conclude that BFR's proponent can deliver a plausible explanation of D2. So, if D1–D3 really do favor premise (2) of the TCA over BFR, that must be because BFR can't honor D3.

But proponents of BFR can accommodate D3. BFR entails that you make a tough choice only if (prior to the choice) you believe you're free to perform multiple different acts. Now, every party to the debate over compatibilism (about metaphysical freedom and determinism) will—or at least should—agree that compatibilism is not *obviously* false.[26] (Indeed, it may well be that most philosophers think freedom and determinism are compatible.)[27] So, since the freedom to perform multiple different acts is not *obviously* inconsistent with determinism, BFR allows that someone who believes determinism could nevertheless make tough choices without thereby holding explicitly inconsistent beliefs. Perhaps somewhat surprisingly, BFR can honor D3.

Of course, BFR *does* entail that someone who makes tough choices while believing she lacks freedom holds explicitly inconsistent beliefs. So BFR entails that a maker of tough choices who is also a *hard* determinist—i.e., someone who believes incompatibilism *plus* determinism—has explicitly inconsistent beliefs. But the thesis that a hard determinist could make tough choices without holding explicitly contradictory beliefs is (to put it mildly) highly questionable.[28] Unlike D3, then, the claim that a hard determinist **(p.53)** could make tough choices without explicitly contradicting himself is much too controversial to count as anything like an adequacy constraint on theorizing about the epistemology of making tough choices.

The upshot of the last several paragraphs is that BFR explains D1–D3 at least as well as does premise (2) of the TCA. My defense of the two questionable steps in my objection to the TCA is now complete. My overall conclusion, then, is that the best attempt to show that being omniprescient rules out acting in certain ways—what we've called the Tough Choices Argument—fails due to its dialectical dependence on the alleged incompatibility of omniprescience and freedom. Contrary to what proponents of arguments like the TCA have claimed, such arguments do not constitute novel problems for the conception of God embraced by traditional theists.[29]

References

Bibliography references:

Adams, R. (1972), "Must God Create the Best?" *Philosophical Review* 81: 317–32.

Carlson, E. (2000), "Incompatibilism and the Transfer of Power Necessity," *Noûs* 34: 277–90.

Coffman, E. J., and Warfield, T. (2005), "Deliberation and Metaphysical Freedom," *Midwest Studies in Philosophy* 29: 25–44.

Finch, A. and Warfield, T. (1998), "The *Mind* Argument and Libertarianism," *Mind* 107: 515–28.

Ginet, C. (1962), "Can the Will be Caused?" *Philosophical Review* 71: 49–55.

Hampshire, S. and Hart, H. L. A. (1958), "Decision, Intention, and Certainty," *Mind* 67: 1–12.

Hunt, D. (1996), "The Compatibility of Omniscience and Intentional Action: A Reply to Tomis Kapitan," *Religious Studies* 32: 49–60.

—— (1997), "Two Problems with Knowing the Future," *American Philosophical Quarterly* 34: 273–85.

Kapitan, T. (1984), "Can God Make Up His Mind?" *International Journal for Philosophy of Religion* 15: 37–47.

—— (1986), "Deliberation and the Presumption of Open Alternatives," *Philosophical Quarterly* 36: 230–51.

(p.54) Kapitan, T. (1990), "Action, Uncertainty, and Divine Impotence," *Analysis* 50: 127–33.

—— (1991), "Agency and Omniscience," *Religious Studies* 27: 105–20.

—— (1994), “The Incompatibility of Omniscience and Intentional Action: A Reply to David P. Hunt,” *Religious Studies* 30: 55–66.

—— (1997), “Acting and the Open Future: A Brief Rejoinder to David Hunt,” *Religious Studies* 33: 287–92.

La Croix, R. (1976), “Omniprescience and Divine Determinism,” *Religious Studies* 12: 365–81.

Mele, A. (2003), *Motivation and Agency* (New York: Oxford University Press).

Merricks, T. (2009), “Truth and Freedom,” *Philosophical Review* 188: 29–57.

Nahmias, E. (2006), “Close Calls and the Confident Agent: Free Will, Deliberation, and Alternative Possibilities,” *Philosophical Studies* 131: 627–67.

Nelkin, D. (2004), “Deliberative Alternatives,” *Philosophical Topics* 32: 215–40.

Pereboom, D. (2008), “A Compatibilist Account of the Epistemic Conditions on Rational Deliberation,” *The Journal of Ethics* 12: 287–306.

Pettit, P. (1989), “Determinism with Deliberation,” *Analysis* 49: 42–4.

Plantinga, A. (1974), *The Nature of Necessity* (Oxford: Oxford University Press).

Quinn, P. (1978), “Divine Foreknowledge and Divine Freedom,” *International Journal for Philosophy of Religion* 9: 219–40.

Rowe, W. (2004), *Can God be Free?* (Oxford: Oxford University Press).

Senor, T. (2008), “Defending Divine Freedom,” *Oxford Studies in Philosophy of Religion* 1: 168–95.

van Inwagen, P. (1983), *An Essay on Free Will* (Oxford: Oxford University Press).

—— (1998), “The Mystery of Metaphysical Freedom,” in P. van Inwagen and D. Zimmerman (eds.), *Metaphysics: The Big Questions* (Malden, MA: Blackwell).

—— (2004), “Van Inwagen on Free Will,” in J. Campbell, M. O'Rourke and D. Shier (eds.), *Freedom and Determinism* (Cambridge, MA: The MIT Press).

Wainwright, W. (1996), “Jonathan Edwards, William Rowe, and the Necessity of Creation,” in J. Jordan and D. Howard-Snyder (eds.), *Faith, Reason, and Rationality* (Lanham, MD: Rowman and Littlefield).

Waller, B. (1985), “Deliberating about the Inevitable,” *Analysis* 45: 48–52.

Warfield, T. (1997), “Divine Foreknowledge and Human Freedom are Compatible,” *Noûs* 31: 80–6.

—— (2008), “Metaphysical Compatibilism's Appropriation of Frankfurt,” *Oxford Studies in Metaphysics* 3: 283–95.

Notes:

(1) La Croix (1976) and Quinn (1978) are early, influential contributions to contemporary work on such issues.

(2) What we'll call “tough choices” result from what Peter van Inwagen (2004: 217) has labeled “serious deliberation” and what Eddy Nahmias (2006: 631) has dubbed “close calls.” Writes van Inwagen: “[S]erious deliberation…[occurs when] one is choosing between alternatives and it does not seem to one (once all the purely factual questions have been settled) that the reasons that favor either alternative are clearly the stronger.” And writes Nahmias: “Let us call these moments of (in)decision, which leave the agent…with nearly equally compelling alternatives for choice, ‘close calls.’ ”

(3) Kapitan's development and defense of this argument spans several papers. See Kapitan (1984, 1990, 1991, 1994, 1997).

(4) See Hunt (1996, 1997).

(5) Following Plantinga (1974), say that a possible world W is feasible iff God has it within his power to “actualize” (make actual) W. For variety, I'll use “feasible” and “creatable” interchangeably.

(6) For helpful discussion and defense of this and other similar closure principles for “relative unfreedom,” see (e.g.) Finch and Warfield (1998) and Carlson (2000).

(7) For dissent, see Adams (1972) and Wainwright (1996). For replies, see Rowe (2004) and Senor (2008).

(8) For similar argumentation, see Rowe (2004).

(9) Cf. Senor (2008: 186–94).

(10) ***Objection***: Suppose it's a brute contingent fact that God prefers a certain combination of values, V, to all other value combinations for possible worlds. Suppose further that among the incommensurable feasible worlds (or the worlds tied for best feasible world, or the infinite hierarchy of better and better feasible worlds) there's exactly *one* that instantiates V. Then it may be there's a creatable world such that God knows his reasons conclusively favor actualizing *it*. But this means the claim to which this note is appended doesn't follow from the reason given for it (viz., that one of the three indicated options holds). ***Reply:*** First, assuming that God must do anything his reasons conclusively favor, the objector's supposition implies (via the closure principle for “relative unfreedom”) that God could be unfree with respect to which feasible world is actual. Since I think (with most traditional theists, so far as I can tell) God must be free with respect to which creatable world is actual, I deny that the objector's supposition is possible. But second (and more importantly), the overall argument in which QS2 figures still prevails (with slight modification) even if the objector's supposition is possible. Recall the argument's target conclusion: *Most traditional theists should think that God makes some tough choices.* If the objector's supposition is possible, what follows from the claim that one of the three relevant options holds is this: *either* each creatable world is such that God knows his reasons don't conclusively favor actualizing it *or* God isn't free with respect to which creatable world is actual. Since most traditional theists think God *is* free with respect to which feasible world is actual, most of them should conclude that each creatable world is such that God knows his reasons don't conclusively favor actualizing *it*—and so, that when God chooses which feasible world to actualize, he makes (what we're calling) a tough choice. (Thanks to an anonymous referee for raising the objection considered in this note.)

(11) Notably, (2) could—without substantive change to the TCA—be weakened so that it requires only that a maker of tough choices be *disposed to believe* about each of her considered courses of action that she may take it.

(12) See, e.g., Kapitan (1997: 287–9).

(13) Mele (2003: 200–1).

(14) Mele (2003: 201).

(15) Nelkin (2004: 221).

(16) Ginet (1962: 50–1).

(17) Hunt (1997: 282).

---

(18) For articulation, argument for, and defense of something like BFR, see van Inwagen (1983, 1998, 2004) and Coffman and Warfield (2005). Notably, BFR could—without substantive change to my objection to the TCA—be weakened so that it requires only that a maker of tough choices be *disposed to believe* about each of her considered courses of action that she's (metaphysically) free to take it.

(19) For compelling defenses of the compatibility of omniprescience and freedom, see Warfield (1997) and Merricks (2009).

(20) Cf. La Croix (1976: 365) and Kapitan (1991: 107).

(21) One could also challenge the indicated conditional with the claim that knowing you'll perform A entails intending to A. Combining this claim with the thesis that you're free only if you sometimes gain intentions would lend support to the claim that there can't be a free omniprescient person. But (as we saw above) knowing you'll perform A does not entail that you already intend to A.

(22) I invite readers to amplify this prima facie case for (2) as they see fit in order to determine whether a suitably modified version of the following argument succeeds.

(23) For helpful discussion of D1 that cites both historical and contemporary proponents of such claims, see van Inwagen (1983), Kapitan (1986), and Nelkin (2004).

(24) See Kapitan (1986, 1991, 1997).

(25) For endorsement of such claims, see Waller (1985), Kapitan (1986), Pettit (1989), Nelkin (2004), and Pereboom (2008).

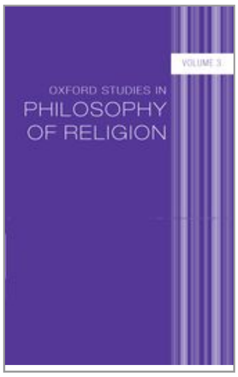
(26) Cf. van Inwagen (1983: 18): “I shall not assume that either of [incompatibilism and compatibilism] is prima facie right. I shall treat them as philosophical theses of equal initial plausibility, and this, it seems to me, is the only reasonable way to approach the Compatibility Problem.”

(27) Cf. Warfield (2008: 283).

(28) For arguments for, and defense of, the thesis that any such *hard* determinist holds inconsistent beliefs, see, e.g., van Inwagen (1983) as well as Coffman and Warfield (2005).

---

(29) Previous versions of this paper were presented at the University of Notre Dame and the University of Oklahoma. Thanks to those audiences for helpful feedback. Special thanks to Andrew Bailey, Nathan Ballantyne, Paddy Blanchette, Monte Cook, Steve Ellis, Reinaldo Elugardo, Thomas Flint, Jennifer Jensen, Jonathan Kvanvig, Matthew Lee, Samuel Newlands, Wayne Riggs, Lee Shepski, Jeff Speaks, Chris Swoyer, Ted Warfield, Linda Zagzebski, and some anonymous referees.



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

Darwin, God, and Chance *

Phil Dowe (Contributor Webpage)

DOI:10.1093/acprof:oso/9780199603213.003.0004

Abstract and Keywords

This chapter considers the options for a person who believes in God's providence, prayer and also free will, to accept Natural Selection. It argues that it is possible. The chapter focuses on the idea that Natural Selection entails that we are here by chance not design. It distinguishes several important senses of chance, then canvases a number of models of divine action, neo-Leibnizianism, neo-Molinism, Occasionalism, and an ‘immediate action’ model. The chapter argues that the latter is fit to the task if one adopts certain views about causation.

Keywords: Providence, Natural Selection, chance, causation, free will

I. Introduction: Darwin Was Wrong

Does Darwin's theory of natural selection entail that we are here by chance? Or at least, would it be unreasonable to deny that we are here by chance even if one thought we are here by natural selection? Darwin thought so. In *The Variation of Animals and Plants under Domestication* he gives an example of a builder who wants to build a house out of stone and goes to the bottom of a cliff to find pieces of stone for this purpose. Some fragments of stone are wedge-shaped for arches, others are elongated for lintels, flat for the roof and so on. Darwin asks:

> [C]an it be reasonably maintained that the creator intentionally ordered, if we use the words in any ordinary sense, that certain fragments of rock should assume certain shapes so that the builder might erect his edifice?[1]

So, too, for natural selection. Could it be reasonably maintained that the creator intentionally ordered variations so that they might be selected for in their particular environment? Darwin's answer is “no.”

I think Darwin was wrong. I do agree that we are here by natural selection, and I agree there are no objective purposes of this sort; however, like many others I don't agree that natural selection and God's detailed design are incompatible. Unfortunately, in this context not much attention has been given to *how*, according to this “compatibilism,”[2] God brings about things that also occur by natural selection. Most compatibilists believe in natural selection and design. Peter van Inwagen believes in design but not **(p.56)** natural selection, nevertheless he argues that they are compatible.[3] I believe in natural selection but not design, nevertheless I too will argue that they are compatible. My argument differs in detail, and is more general, than van Inwagen's.

One thing that is striking about Darwin's stone house example is that, whether or not it can be *reasonably maintained*, I think there is no doubt that there are many believers who would at least *maintain* that the stones were there because of God. Place the example in a context where the builder desperately needs to build the house, and quickly, and I think you have a situation where believers typically pray for God's help, and indeed thank God when they find the right-sized stones for their task. In other words, such believers do believe that the creator intentionally ordered that certain fragments of rock should assume certain shapes. If I am right that natural selection and God's detailed design are compatible, one might hope that showing so could help convince at least some of those believers who reject natural selection to rethink their position. If this is the higher goal, then it will not suffice to demonstrate the compatibility of natural selection and God's detailed design should that demonstration seem to threaten the idea that God looks after his creatures, and answers prayers. This also raises questions about free will and the causes of evil. These issues all concern if and how God causes things.

II. Chance, Coincidence, and Accident

The word “chance” has many meanings. In one sense of chance it is indeed true that if we are here by chance then we are not designed. But in that sense of chance, natural selection does not entail we are here by chance. In another sense of “chance,” natural selection does indeed entail that we are here by chance. Chance variations are selected for because they are advantageous in their environment. But in that sense of chance, natural selection does not rule out God's having designed us. Let's distinguish three senses of “chance”:

1. *Objective single-case chance*. If there are events that lack a sufficient cause, then we have objective single-case chance. The laws of nature may bestow a probability on an event's occurrence, but that probability is less than one. Darwin did not believe that variations occur by chance in this sense. Like most in the nineteenth century, Darwin believed in Laplacean **(p.57)** determinism, that every event has a sufficient cause. A complete description of the world at a time, together with the laws of nature, fixes all future events.

2. *Coincidence*. John and Mary bump into each other in a bar *by chance*. There may well be a completely deterministic chain of events that led to John's going to that bar, and a completely deterministic chain of events that led to Mary's going to that bar. The co-occurrence is still by chance in the sense that there is no causal explanation for why they are both there together: John is not there because Mary is there, Mary is not there because John is there, and there's no common cause for them both being there. Two co-occurring events constitute a coincidence, let's say, if the chain of causes leading to each are independent.[4] (The meeting would not be a coincidence if John went there because he knew Mary would be there.)

In his influential book *The Direction of Time* Hans Reichenbach sets out the Principle of the Common Cause, a precise way to decide whether a co-occurrence is a coincidence or whether it has a causal explanation. (We'll leave aside a host of worries about inferring causes from probabilities.) Some “coincidences” seem too improbable to really be coincidences, Reichenbach observed, because they occur together too often to be put down to “chance.” The co-occurrence of A and B is a coincidence just if:

$$P(A.B) = P(A)P(B)$$

(1)

where the probabilities are understood as frequencies. Suppose Mary is in financial difficulty, and prays for help, and that very morning an envelope containing a significant sum of money appears anonymously under her door. Is this a coincidence? According to Mary it is no coincidence—she says God answered her prayer—but her friends say it is just a coincidence. Reichenbach's analysis can settle this. Suppose it is quite common that people in financial trouble pray for help (A), suppose P(A) = 1/100, i.e. around the world every day 1 in 100 people in the context of financial trouble are praying for help. Suppose finding a significant sum of money under your door (B) is not so common, suppose P(B) = 1/100,000, i.e. every day 1 in 100,000 people around the world find significant sums of money under their door. Then, by “chance” we would expect the co-occurrence to occur 1/100 x 1/100,000 of the time, i.e. every day 1 in 100,000,000 people around the world will find their prayer followed by the appearance of money. This shows that Mary's case is just a coincidence (in the absence of other relevant factors). On the other hand, in cases where (1) does not hold, **(p.58)** the Principle of the Common Cause says that there must be a causal explanation, either a direct causal link between A and B or a common cause of A and B.

On Darwin's theory of natural selection variations occur by chance in the sense of a coincidence. It is a coincidence that a variation occurs in a favorable environment (compare van Inwagen 2003, 354, 361). That it does not occur because the environment is favorable is an essential part of the natural selection explanation of why we are here. Evolution by natural selection, Darwin said:

> absolutely depends on what we in our ignorance call spontaneous or accidental variability…the variations of each creature are determined by fixed and immutable laws; but these bear no relation to the living structure which is slowly built up through the power of selection…[5]

3. *Accident*. If I find treasure in a ditch (Aristotle's example) *by chance*, that means there is no teleological connection between my digging there and the finding the treasure. I did not dig in order to find the treasure.[6] It's often asserted that variations are accidents: they do not occur in order that they be selected. No one intended us to be here. But if God plans or deliberately guides evolution, then at least some variations are no accident. The incompatibilist argument can be put in these terms: natural selection entails coincidences, and these must be an accident so they cannot be designed. The enthymeme is that nothing that is a coincidence can fail to be an accident. Not everyone agrees with this enthymeme. Jung's theory of synchronicity is a denial of this.[7] I will not follow Jung, but I will assert that the enthymeme is true in one sense but false in another.

III. Neo-Leibnizianism

Consider Leibniz's picture of the world. For every feature of the world there must be a reason why things are that way rather than some other way. In creating the world, God considers all the possible ways the world could be, and chooses one because it is the very best. God then creates that world and decrees that it develops deterministically. So every small feature of the world is there because God deemed that things are better that way. However, for **(p.59)** Leibniz there is a sense in which the unfolding world is an emanation of God minute by minute.[8]

Let's take a *neo-Leibnizian* to hold that God designed and created everything according to a minutely detailed plan and that this world develops by Laplacean determinism. Everything has a reason, including coincidences, but unlike Leibniz the neo-Leibnizian is not asserting that the unfolding world is an emanation of God. In the neo-Leibnizian world there is no single case-objective chance, and there are no accidents or coincidences. God has deliberately arranged things so that John meets Mary at the bar, and that I find the treasure.

So can a neo-Leibnizian believe in natural selection? Yes, God arranges things so that the right pieces of stone fall to the bottom of the cliff in order that the builder finds the right pieces for building the house. Likewise, God arranges things so that each variation occurs at the right time and place in order that natural selection operates, in order that (among other things) we are here.

What happens to the claim that variations occur by chance, on the neo-Leibnizian picture (i.e. it's a coincidence that a favorable variation occurs in a specific environment)? That the variation occurs in a favorable environment is no coincidence—God is the common cause. But we can say this is a *physical coincidence*; i.e. there is no physical causal explanation where a physical causal explanation is one that appeals only to physical (or natural) causes[9] (which do not include God). Reichenbach's equation (1) would still hold. So we can't claim that natural selection entails that we are here by chance and hence there is no God who intentionally arranged for us to be here. The sense in which natural selection entails that we are here by chance (physical coincidence) is not a sense that entails there is no God who intentionally arranged for us to be here. On the other hand, if it is an accident that we are here, then indeed there is no God who intentionally arranged for us to be here. But natural selection does not entail that it is an accident that we are here. Intelligent design and natural selection are compatible.

Neo-Leibnizianism may be acceptable to some contemporary believers who currently reject evolution, and who also believe in predestination: God chooses in advance who will be saved, and brings it about that they come to believe. Not all believers agree, holding instead that God allows free will to choose to believe or not. Leibniz himself had a compatibilist view of free will, but this may not be acceptable to some believers, since it still has God **(p.60)** causing the choice to believe. Anti-determinist believers typically are libertarians: (on one version) free actions are caused by free choices, and those free choices are uncaused.[10] It may also be objected that science (i.e. quantum physics) shows that the world isn't in fact deterministic in the manner of neo-Leibnizianism, so this can't be how God does it.

IV. Neo-Molinism

Suppose then that determinism is false. Suppose there are uncaused causes—libertarian free choices—originating let's say for simplicity in Cartesian souls. There are also ubiquitous chancey physical events. Then we must give up the Leibnizian/Laplacian story where knowledge of the complete state of the universe at a time together with knowledge of the laws enables prediction of all future events.

According to most contemporary accounts of causation (for exceptions to this, see below), God might still cause everything. A “cause” according to this view is any causal factor, and typically an event has many causal factors. If a cause is anything on which the effect counterfactually depends,[11] then since whether one event counterfactually depends on another is determined by background conditions, the effect will counterfactually depend on each of these background conditions taken separately. If a cause is anything that raises the probability of the effect.[12] then since whether one event raises the probability of another also depends on background conditions, each of these background conditions taken separately will also in general raise the probability of the effect. So, when a Cartesian uncaused free choice has a physical effect, whether it does so will depend on physical background conditions each of which will also count as a cause of the physical effect. Also, chancey physical events are caused (probabilistic causation) on the two accounts just mentioned.[13] And even if one holds a theory of causation that rules out probabilistic causation,[14] a chance event will have effects—and an effect being caused depends on background conditions some of which may not themselves be chance events, and those background conditions will count as causes of the effect, and will themselves be caused. So, assuming the transitivity of causation, even in such an indeterministic world God could be the cause of every physical event even though he does not cause the uncaused choice. The array of physical events in the world then would be **(p.61)** the co-production of God's act of creation and free choices and other chance events.

However, if God knows the future only by self-knowledge of his own creative act, he would know only an array of possible outcomes, a “multiverse” if you like, but would not know what the actual array of physical events will be. Thus, although he may be the cause of all physical events, if he doesn't know what the effects will be of his act of creation, God could not be said to bring them about for a purpose. It would be an accidental world.

According to Molinism God foresees the outcomes of free human choices, and accounts for these in creating the world. To do so God has “middle knowledge,” knowledge of counterfactuals of the form “were substance s to be created s would choose a in circumstances C.”[15] To get “neo-Molinism” add to this genuine chance events, such that according to the laws state S_1 of substance s gives state S_2 a probability p in circumstances C. And God has middle knowledge of the form “were s to be in state S_1 in circumstances C, state S_2 would follow.” (This possibility was suggested to me by David Lewis, July 1999.)

So can a neo-Molinist believe in natural selection? Yes, God arranges the laws and initial conditions, seeing the outcomes of chance events, in order to bring about each physical event according to his plan. So again, God arranges things so that the right pieces of stone fall to the bottom of the cliff in order that the builder finds the right pieces for building the house, and that each variation occurs at the right time and place in order that natural selection operates, in order that we are here.

Would contemporary intelligent designers be happy with either the neo-Leibnizian or the neo-Molinist God who arranged for us to be here via natural selection? Some may object to a perceived deism,[16] with a God that is much more remote than the God they take to be the object of their experience. When the believing builder prays that he might find the right stones for his house when he goes to the cliff, and on finding them, gives thanks to God, there's an implicit sense that God is acting immediately, indeed, in response to the prayer. Leibniz and many other subscribers to “concurrence” throughout history deny that their picture leads to deism. For Leibniz, God not only creates the world but also sustains it by holding in being at all moments all substances and their causal powers. And God does “respond” to prayer by seeing in advance what the prayer will be and arranging the “answer” in advance. But saying that God knew in advance what the believer **(p.62)** would ask for and has prepared the answer in advance may not cut any ice with the believer. It seems to make God's interest in our affairs temporally remote. Believers hold that God literally answers prayer.

V. Occasionalism

According to occasionalism,[17] God immediately brings about everything that occurs, in accordance with a plan and in a manner according to which regularities in nature can at least sometimes clearly be discerned. Created beings, animate and inanimate, do not actually cause anything. This follows, for example, from Malebranche's notion of a cause as logical necessitation. Only from God's decision to do something would it involve a logical contradiction to suppose that the intended effect did not follow, given that only God is infallible and all-powerful. Occasionalism can allow for God's being the cause of a variation according to a plan as much as neo-Leibnizianism. When the time comes, God intentionally brings about a variation that has certain fitness in its environment, and clearly can do so for the reason that it has that fitness. And when the time comes, God intentionally causes the rocks to fall in a certain shape, according to a plan. And since God clearly could have done otherwise, the occasionalist can say God does so in response to the believer's prayer.

However, one major objection to occasionalism is that it denies humans free will. Leibniz, for example, objected to Malebranche's occasionalism on the grounds that it robbed created things of all real causal power. Further, since neo-Leibnizianism, neo-Molinism, and occasionalism each make God the cause of all physical events, there's the related problem of God being the cause of evil. In the case of neo-Molinism, God does not cause the sinner to choose to sin, but God does participate in causing the effects of that choice. When an innocent child is murdered, both God and the murderer are causes. If there are physical events that are intrinsically evil, then God knowingly causes them. (I won't address the problem of so-called natural evil.)

VI. Chance and Immediate Action

Suppose God creates a world with chancey laws, but as the possibility of a chance event lawfully arises, he either allows them to occur by chance or **(p.63)** causes their outcome by immediate action.[18] Suppose for the moment that there is no such thing as middle knowledge. God can start with at least a partial plan (including, for example, that conscious beings evolve) that goes beyond what is possible based on his knowledge of the array of possible outcomes, because he intends to intervene to preserve the plan. On this view God allows some events to proceed by chance not knowing in advance what the outcome will be. If things are not heading the right way, he can steer the world back to the plan by intervening immediately in other chance events.

These interventions will not count as miracles if miracles are defined as violations of laws of nature. We assume here that God's immediate actions are random with respect to, say, quantum mechanics, so that the frequencies predicted by quantum mechanics are not changed by the array of God's immediate interventions. However, there's no reason why the immediate action story could not be told with miracles as well as or instead of interventions on chance. For simplicity I will focus only on interventions on chance.

According to this picture God sometimes arranges things—perhaps in response to prayer—so that the right pieces of stone fall to the bottom of the cliff in order that the builder finds the right pieces for building the house, and that sufficient variations occur at the right time and place in order that natural selection operates according to his partial plan, in order that we are here. How often God intervenes by immediate causation depends on how detailed his original plan is, and on what chance events occur including what prayers occur. That a certain variation arises in a given environment is no coincidence, however the causal structure may not be one of a common cause—it may be direct causation where God sees the environment and brings about a favourable variation on account of that.

Is it nevertheless a physical coincidence that a favourable variation occurs in a particular environment? In the latter case, it would seem not. If God is an intermediate cause between the environment and the variation, then at least on the probabilistic theory of causation the variation and the environment cannot be statistically independent as in Reichenbach's equation (1) above. If this is right then it would seem the proposal about God's causing variations is empirically falsifiable. It's not clear exactly how that would work, since we have only the vague proposal “God guides evolution.” Under neo-Darwinianism mutations provide the variation on which natural selection works, and the ratio of favourable to unfavourable mutations is **(p.64)** extremely low. God would not so much increase this ratio as change the kinds of mutations occurring, in order to direct selection. At this level of generality it is difficult to see how the correlation would be visible. But it seems in principle falsifiable. To my mind this is a problem, but my response to it is the NMP response—Not My Problem. The believer whom I am addressing here already believes that God literally answers prayer. It adds nothing new to the picture to claim God guides evolution since this believer already has the problem of in-principle falsification.

One could have a neo-Molinist version of the immediate action account. God sets a general plan in train in creating chancey laws and initial conditions. Having middle knowledge of chance outcomes, God also plans to respond with immediate interventions at the right times and places. Then on either version God can be said to guide evolution, to allow for libertarian free will, and to (literally) respond to prayer.

But does God cause evil? It would seem so, as follows. Say there is an intrinsically evil physical event, assuming again that there are such things, caused by an uncaused evil libertarian choice. Say this evil event was not part of God's plan. On the non-Molinist version God doesn't know in advance of the choice that this will happen. First, it must be the case that God, being all-powerful and all-knowing, could have prevented the evil by intervening between the evil choice and the evil event. One can insist nonetheless that God did not thereby cause the evil event by distinguishing allowing from causing. There are accounts of so-called “absence causation” according to which this is so.[19] But second, that God does not cause the evil event on account of the fact that he could have prevented it does not mean God did not cause the evil event. He still causes it as a partial cause together with the evil choice, having provided the background conditions in which the evil choice occurs.

Is there an account of causation according to which this is not so? Yes. According to John Stuart Mill[20] a total cause is the full set of conditions that together form an invariable sequence with the effect, where each of these conditions is a necessary part of the total cause (necessary in order to form an invariable sequence). For Mill, like David Hume, the invariable sequence is just an actual universal regularity, but one could think of the total cause in terms of a nomologically sufficient condition of the effect, if one believed in a stronger view of laws. According to Mill, any of the conditions that together form the total cause can be deemed “a cause” so long as we don't think that any one condition has some objective right to be called a cause **(p.65)** rather than any of the other conditions. In other words, Mill rejects the distinction between causes and conditions. This is the position that I said above is pretty much standard in contemporary philosophy of causation, and that I have assumed so far.[21] We can allow this to stand for physical causation, but introduce a condition on mental causation familiar in the law. Suppose more than one mental event is part of the causal history of a particular effect. Call the proximate, or closest, of these mental causes *the cause* and all others *conditions*. “Closest” can be defined either spatio-temporally or causally. I won't attempt to define the latter because it would depend on exactly which theory of causation one is working with, and in any case we want to take the former option here. Then, in the case of the intrinsically evil physical event caused by an uncaused evil libertarian choice, the evil choice is the cause of the physical event, and God, having provided the background conditions in which the evil choice occurs, is a mere condition, providing God does not intervene after the evil choice occurs.

Note, finally, that this account is available to the compatibilist as much as to the libertarian. My point is simply that the believer who wishes to wash God's hands of this kind of evil can believe in natural selection, as it is a relatively simple matter to provide an account of causation that can do this.

References

Bibliography references:

Darwin, C. (1875), *The Variation of Animals and Plants Under Domestication*, vol. 2., 2nd. edn. (London: John Murray).

Dowe, P. (2000), *Physical Causation* (New York: Cambridge University Press).

—— (2004), “Causes are Physically Connected to their Effects: Why Preventers and Omissions are not Causes,” in C. Hitchcock, (ed.), *Contemporary Debates in Philosophy of Science* (Oxford: Blackwell), 189–96.

—— (2005), *Galileo, Darwin, and Hawking: The Interplay of Science, Reason, and Religion* (Grand Rapids, MI: Eerdmans).

Flint, T. (1998), *Divine Providence: The Molinist Account* (Ithaca, NY: Cornell University Press).

Freddoso, A. (1991), “God's General Concurrence with Secondary Causes: Why Conservation is not Enough,” in James E. Tomberlin (ed.), *Philosophical Perspectives* (Atascadero, CA: Ridgeview Publishing), 5: 553–85.

Hart, H. L. A., and Honoré, T. (1985), *Causation in the Law* (Oxford: Clarendon Press).

(p.66) Hume, D. (1748), *An Enquiry Concerning Human Understanding* (London).

Ide, H. A. (1993), “Aristotle, Metaphysics vi 2–3 and Coincidences,” *Ancient Philosophy* 13: 341–54.

Jung, C. [1952] (1960), “Synchronicity: An Acausal Connecting Principle,” reprinted in H. Read, M. Fordham and G. Adler (eds.) (1960), *The Collected Works of C. G. Jung* (London: Routledge & Kegan Paul), 8: 417–531.

Kvanvig, J. and McCann, H. (1991), “The Occasionalist Proselytizer: A Modified Catechism,” in James E. Tomberlin (ed.), *Philosophical Perspectives* (Atascadero, CA: Ridgeview Publishing), 5: 587–615.

Leibniz, G. W. (1975), “Discourse on Metaphysics,” in G. Parkinson (ed.), *Philosophical Writings* (Totowa: Rowman and Littlefield). (AQ)

Lewis, D. (1973), “Causation,” *Journal of Philosophy* 70: 556–67.

Malebranche, N. (1997), *The Search after Truth: with Elucidations of The Search after Truth*, ed. and trs. T. M. Lennon and P. J. Olscamp (Cambridge: Cambridge University Press).

Menzies, P. (2004), “Difference-making in Context,” in J. Collins, N. Hall, and L. Paul (eds.), *Causation and Counterfactuals* (Cambridge, MA: MIT Press), 139–80.

Mill, J. S. (1843), *A System of Logic* (London: Longman), book III, chapter V.

Molina, Luis de (1988), *On Divine Foreknowledge: Part IV of the Concordia*, trs. A. Freddoso (Ithaca, NY: Cornell University Press).

Owens, D. (1992), *Causes and Coincidences* (Cambridge: Cambridge University Press).

Polkinghorne, J. (1986), *One World: The Interaction of Science and Theology* (London: SPCK).

Reichenbach, H. (1956), *The Direction of Time* (Berkeley, CA: University of California Press).

Stump, E. (2001), “Augustine on Free Will,” in E. Stump and N. Kretzmann (eds.), *The Cambridge Companion to Augustine* (Cambridge: Cambridge University Press), 124–47.

Suppes, P. (1970), *A Probabilistic Theory of Causality* (Amsterdam: North-Holland Publishing Company).

van Inwagen, P. (1988), “The Place of Chance in a World Sustained by God,” in T. Morris (ed.), *Divine and Human Action* (Ithaca, NY: Cornell University Press), 211–35.

—— (2003), “The Compatibility of Darwinism and Design,” in N. Manson (ed.), *God and Design: The Teleological Argument and Modern Science* (London: Routledge), 348–63.

Notes:

(*) I'd like to thank anonymous referees for helpful comments.

(1) Darwin (1875: vol. 2: 427).

(2) The term is from P. van Inwagen (2003: 349).

(3) Van Inwagen (2003).

(4) Owens (1992).

(5) Darwin (1875: vol. 2: 236).

(6) Ide (1993: 341–54).

(7) Jung [1952] (1960: vol. 8, 441).

---

(8) Leibniz [1686] (1975).

(9) Dowe (2000).

(10) Stump (2001: 125–6).

(11) Lewis (1973: 556–67).

(12) Suppes (1970).

(13) See Dowe (2000, ch. 2).

(14) Hume (1748: sec. 7).

(15) Molina (1988); Flint (1998).

(16) van Inwagen (1988: 215, n. 4).

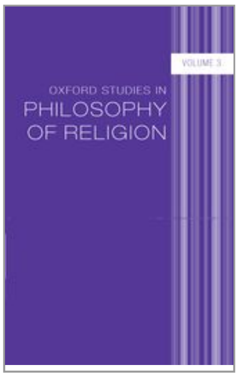
(17) Malebranche (1997); Kvanvig and McCann (1991).

(18) See Polkinghorne (1986: 69); for a critique see Dowe (2005: 184–5) and compare Freddoso (1991).

(19) E.g., Dowe (2004).

(20) Mill (1843, bk. 3, ch. 5).

(21) Exceptions include Hart and Honoré (1995) and Menzies (2004).



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

Intelligent Design and Selective History: Two Sources of Purpose and Plan

Peter J. Graham (Contributor Webpage)

DOI:10.1093/acprof:oso/9780199603213.003.0005

Abstract and Keywords

In Warrant and Proper Function and in Knowledge of God, Alvin Plantinga provides an intelligent design argument for the existence of God. Plantina's key premise is that “naturalism” cannot account for the proper functions of natural traits, but God as intelligent designer can. His argument for this premise relies on a thought experiment involving a Hitler-like madman and his henchmen. This chapter criticizes Platinga's thought experiment through careful attention to both intelligent design and natural selection as sources of functions and standards for normal functioning.

Keywords: functions, God, intelligent design, natural selection, normal functioning, plantinga, proper function

Intelligent design arguments follow a familiar pattern. They first specify a seemingly unobjectionable fact. Next they claim that no natural process could produce it, but a powerful, rational, intelligent designer could. *Voilà*, a powerful, rational, intelligent designer and creator, presumably the God of Abraham, therefore exists.

In *Warrant and Proper Function* (1993) and then again in *Knowledge of God* (2008), Alvin Plantinga gives the following intelligent-design argument.[1] First, we correctly apply the teleological terms “function,” “purpose,” “proper functioning,” “design,” and “design plan” to various parts of living organisms, just as we correctly apply these terms to intentionally designed, created artifacts. Second, naturalism cannot explain these applications; there is no adequate naturalist account of function, purpose, proper functioning, design, and design plan. In particular Plantinga argues they all face a devastating counter-example involving a Hitler-like madman and his henchmen. Third, the supernaturalist can explain the applications. From the supernaturalist perspective “it is easy enough to say what it is for our faculties to be working properly: they are working properly when they are working in the way they were intended to work by the being who designed and created both them and us.”[2] So we possess a “powerful argument” that God exists:

> …suppose…you are convinced…that [there are] such things as proper function, damage, design, dysfunction, and all the rest…. if you think there is no naturalistic **(p.68)** analysis of these notions, what you have is a powerful argument against naturalism. Given the plausible alternatives, what you have, more specifically, is a powerful theistic argument; indeed, what you have is a version of Thomas Aquinas's Fifth Way.[3]

Plantinga offers this argument while developing a proper-function reliabilist theory of warrant. That makes him an intelligent-design reliabilist.

Like most words in English, the word “function” has more than one use or meaning. One use involves *purpose*, where a function of an item is *what the item is for*. Ballpoint pens are for writing; that is their function or *purpose*. The function of the heart is to pump blood; that is what it is for. Another use involves capacity or ability, where a function of an item involves what it *can* do; functions in this sense are causal role *capacities* or *dispositions*. A third use involves how an item with connecting parts *works* or *operates*; this use involves how the item functions. Mechanics know how cars work; they know automotive design. Physicians study how our bodies operate.

The present debate concerns the first and third uses of “function”: purpose and plan. I contend that there are at least two sources of purposes and plans: intelligent design and selection over variants in a population over generations. The first source is familiar to everyday life. The second source is best understood with a layman's grasp of evolution by natural selection, but applies to all sorts of items, not just biological traits.

In the first section I explicate deliberate, conscious, intelligent design as one kind or source of purposes and plans. In the second I explicate selective history as another. In the third I state and criticize Plantinga's counter-example. Plantinga's case does not withstand a thorough understanding of the two sources.

The discussion concerns the *kinds* or *sources* of purposes and plans, not the kinds of things that *have* purposes and plans. Any item that undergoes selection could have purposes and plans from selection. And any item that can be created by an intelligent designer could have a purpose or plan. So artifacts could have purposes and plans from artificial selection, and if Plantinga is right, the traits of living organisms have their purposes and plans by virtue of intelligent design. The intelligent design/selection distinction thus differs from the artifact/organism distinction.

I. Purposes and Plans From Intelligent Design

I'll begin with purposes and plans from conscious, rational, intentional design. Plantinga relies upon this source in his account of warrant. I will **(p.69)** first state his account to get the relevant teleological terms on the table, reserving discussion of their use in epistemology for another occasion:

> a belief B has warrant for you if and only if (1) the cognitive faculties in the production of B are *functioning properly*…; (2) your *cognitive environment* is sufficiently similar to the one for which your cognitive faculties are designed; (3) the triple of the *design plan* governing the production of the belief in question involves, as *purpose or function*, the production of true beliefs…; and (4) the design plan is a *good one*…[emphasis added].[4]

So how should we understand *design plans*, *proper functioning*, *good design*, and *cognitive environments* from intelligent design?

The design plan for a device includes the *specifications* for how it is *supposed* to work. The design plan is the item's *blueprint*.[5] It's a normative notion for correct or proper operation of the item. To work or operate properly *just is* to work or function according to the design plan.[6] Plantinga uses the phrase “proper function” as short for proper functioning, and by *proper* functioning he means functioning (operating, working) *according to the design plan*.

But we don't simply build devices to work or operate a particular way (though sometimes we might, just to discover what it would do). We also build devices to produce a certain effect. We want them to work a certain way to do something, to solve a problem or produce a desired outcome. In so doing we invest them with *purposes* or *functions*.

A *good* design plan is *effective*. Plantinga is well aware—as we all are—that things do not always go according to plan: some design plans are not very good; they do not produce their desired effect, even when followed to the letter. Some devices work just fine; others do not work at all.

The possibility of *poor* design drives a wedge between proper functioning—operating according to the design plan—and function fulfillment—fulfilling the purpose or function of the item. You can work “properly” without doing what you are supposed to do for you are poorly designed:

> Another important distinction is that between what a thing is designed to do (its purpose, say) and how it is designed to accomplish that purpose (its design, we might say)…a thing works properly when it works in accordance with its design plan…I aim to make a refrigerator that will keep its contents at a constant temperature of 33F; through incompetence and inattention, I fail; as a result of a lamentably inferior design, its internal temperature varies between 70F and 85F. Then the thing **(p.70)** works properly in one way: it works just as it was designed to work. But in another way it works badly: it does not do what it was designed to do.[7]

So an item can function properly (work according to specifications) without fulfilling its function (without producing the effect it was made to produce).

Additionally an item can function properly without fulfilling its function for it is not in normal conditions. When you take your car to the shop and the mechanic puts it up on the lift, it can work exactly as it was designed—your car is in mint condition—but it is not hurtling you down the highway at 65 miles per hour.

This takes us to our last question: what counts as your “cognitive” environment? “Cognitive” environments are simply normal conditions for belief-forming processes. So in general what makes conditions normal?

The answer is straightforward. Normal conditions are those conditions the designer had in mind, either intensionally (the designer had a description of the circumstances in mind) or extensionally (the designer simply pointed at some circumstances or placed the device in those circumstances), for the use of the device.

The overall picture is this. “Intelligent designers” find themselves with the desire or motivation to create an item that will produce a certain effect. That effect is the *function* or *purpose* of the item. The designer will come up with a *plan* for the device, the blueprint for how the device will work so as to achieve the intended purpose. Once constructed, the hope is that the device will turn out to be a *good* one, that it will achieve the intended effect when working according to the design plan. And all of this is relative to intended circumstances for the use of the device. When all goes well, when functioning normally in normal conditions, it will fulfill its function. We can now explain three further features of devices with purposes and plans from intelligent design: the distinction between functions and “accidental” side effects; the plurality of functions; and change in function.

Devices produce a number of effects. The intended effects are functions; the non-intended effects are “accidental” side effects. The function of your watch is to tell time. The noise it makes, or the mark it leaves on your wrist, is but an “accidental” side effect. Your watch was not designed in order to make noise or to leave a mark on your wrist, even though it does both.

A device may also be designed to produce a number of effects. A knife is for cutting all sorts of things. It may also be used as a weapon, a gift, **(p.71)** or a ceremonial device. It may even be a work of art or a status symbol. Items intentionally designed can be designed to produce any number of effects.

We can also *redesign* devices for *new* purposes. The items thereby acquire new purposes and design plans. Such alteration can lead to unproblematic cases where an item functions properly according to the *new* plan, but not according to the *old* one:

> a thing can obviously acquire a *new* purpose and a *new* design plan. I construct my refrigerator and then make a minor adjustment; now the refrigerator works and is intended to work slightly differently. *It has acquired a new design plan. Relative to its old design plan, it is not functioning properly; relative to the new, it is.* Most drastically, I turn my refrigerator into a food warmer by reversing a crucial circuit; then when it functions *properly* according to its *new* design plan, it works *badly* indeed according to the *old*. So suppose our cognitive faculties are redesigned by Alpha Centaurians, whose aims here are aesthetic rather than epistemic. Then it might be that our faculties work in accords with the *new* design plan [but not with] the *old* design plan…[emphasis added].[8]

So if we are devices created by an intelligent designer, it is possible that our designer, or some other intelligent being, should be able to alter our cognitive faculties so that they acquire *new* functions and *new* design plans. Relative to their *old* design plan, they are *malfunctioning*. Relative to the *new*, they are functioning *properly*.[9]

(p.72) II. Purposes and Plans From Selective Histories

To get a handle on selection as the source of purpose and plan, a layman's understanding of evolution by natural selection should help.

Take a population of beetles that *vary* in color. Some are green and some are brown. Suppose they live in a leafy green environment where birds prey on them. Against a green background, the brown beetles are easy pickings. The green beetles, on the other hand, are well camouflaged. Having green coloration is clearly more adaptive than being brown. The green beetles are apt to survive longer and have more offspring than the brown beetles. Surface coloration has *consequences* for survival and reproduction. And assuming that coloration is a *heritable* trait, the number of green beetles will increase in the population over time, perhaps to the point where all the beetles are green, and occasional mutations leading to different-colored beetles are weeded out. Selection leads to a trait clearly adapted to the environment.[10]

Natural selection involves at least three factors. The first is *heredity*. Descendants in a population have the trait because their ancestors in the population had the trait. The second is *variation* among the heritable trait. The third factor is *consequences*. The differences in heritable traits must have consequences for how likely it is that organisms possessing the trait will live long enough to reproduce successfully. The variations must either enhance or diminish relative fitness. A trait that contributes to fitness is *adaptive*, and so likely to be passed on. Over time nature weeds out the traits that diminish relative fitness, and preserves the traits that enhance relative fitness. Traits that result are *adaptations*.[11]

Adaptations thus result from a history of selection for particular adaptive effects. The heart was selected for pumping blood but not for making noise. **(p.73)** There was selection *of* noise-making hearts but not selection *for* noise-makers.[12]

How is natural selection a source of purposes and plans? The etiological theory of functions provides answers.[13]

Advocates of the etiological theory of functions differ over its correct formulation.[14] Here is a formulation that should do. I'll make one revision later:

> A token *t* of type *T* has *F* as an etiological function *iff*:
>
> 1. *t* is a reproduction or copy of previous tokens of type *T*.

2. ancestors of *t* (earlier members in the lineage) produced *F*.
3. there was selection for *T* because ancestors produced *F*.

On the etiological theory, the function of a token *t* of type *T* is thus the selected effect of previous tokens of the type.

(p.74) Take our beetle. Providing camouflage is a selected effect of being green. Camouflage from predators is thus its function; that is what it's for; that is its purpose. It is sufficient on the etiological theory that if you are an adaptation for *F*, producing *F* is among your functions.

But it is not necessary. Whereas “adaptation” is a technical term in biology, the etiological theory covers more territory. Copying or reproducing need not mean biological reproduction—any means of copying or reproducing will do. A lineage of ancestors and descendants need not be biologically related. Any effect that contributes to the differential copying and reproducing of variants of tokens of a type thus leads to functions.

This more general or abstract way of understanding functions means that all sorts of things can have etiological functions. Not just skin-pigmentation patterns, hearts, teeth, eyes, and mating displays, but words, customs, and even artifacts, among other things. In artificial selection we humans are the external selection pressure. But the mechanics of natural selection and artificial selection are the same. And artificial selection need not even be restricted to living organisms. Artifacts, too, get copied and selected for certain of their effects. A belt buckle might get mass produced because it effectively holds pants up at an affordable price. Ruth Millikan's overall philosophical project, in fact, is to use the etiological notion of function to examine different kinds of meanings for words and other linguistic devices, and the origin and content of mental representations and the behaviors they produce.[15]

For items with etiological functions, accidental side effects are then just those effects of the item that played no role in its selection.

Many items with etiological functions have more than one function. The tongue, for instance, has a number of functions; it was selected for a **(p.75)** variety of adaptive effects. The tongue is for talking, tasting, chewing, and swallowing.

For items with etiological functions, what determines normal functioning and what makes conditions normal?

The answer is straightforward. For any item with an etiological function, there is a historical explanation for how the item came to have that function: by having a certain effect while operating in a certain way in circumstances of a certain kind. A muscle in an organism's chest pumps by beating regularly. In turn it is connected in a systematic way with other parts of the organism, embedded in a certain kind of habitat. If pumping blood explains, in part, why the muscle gets reproduced, then it comes to have pumping blood as a function. For any item with an etiological function, there is a historical explanation that explains how the item functions, by working a certain way in certain conditions.[16]

Notice the way the item functioned (operated, worked) when it selected for producing the effect that became its function. That way of functioning is normal functioning. Notice the circumstances in which the item underwent selection. Those conditions and conditions of similar type are normal conditions.[17]

The way the item worked or operated fixes its design plan, for its design plan is just how it is supposed to work or operate; normal functioning is proper functioning.

There is nothing mysterious or objectionably circular in this way of explaining normal conditions and normal functioning. Indeed, it is all rather straightforward. Just look at the history of the item, the effects **(p.76)** that help explain why it gets replicated, how it worked so as to produce these effects, and where it all happened. The historical explanation spells out how it worked, where it did it, and why it was selected. Functions, normal functioning, and normal conditions all fall out. On the etiological notion, functions, normal functioning, and normal conditions are all explanatorily interrelated. Here is a case of interlocking necessary conditions.[18]

An item with an etiological function can surely fail to fulfill its function when not in normal conditions. None of us can see a thing in total darkness. An item with an etiological function may also fail to function normally. A heart may be malformed for a number of reasons. So just as an item with an intentionally assigned function may “malfunction” for failure to fulfill its function (purpose) or failure to function normally (operate as it should), so, too, may an item with an etiological function.

Etiological functions can be lost. The human appendix no longer has a function, though it once aided in digestion. The eyes of some cave-dwelling creatures are no longer for sight but simply reminders of their evolutionary past. Etiological functions are lost when, over time, the trait ceases to produce the effect for which it was selected. So our definition of function needs to be altered somewhat. If, in recent history (which may only be recent in evolutionary time), the item no longer produces the effect it was selected for, then the item loses producing that effect as a function. The appendix is no longer for assisting digestion, and the eyes of fish in lightless caves are no longer for vision.

Etiological functions can also change. An item once selected for a certain effect can undergo new selection pressures where it is selected for a different effect. The item can thereby lose its old function and acquire a new one. Or it may retain its old function while acquiring a new one. Feathers were originally for temperature regulation, like the way our hair follicles assist in sweating. But feathers also came to assist flight.

Taking these facts to heart, Peter Godfrey-Smith[19] suggests that we add a fourth clause to our definition:

> A token *t* of type *T* has *F* as an etiological function *iff*: **(p.77)**
>
> 1. *t* is a reproduction or copy of previous tokens of type *T*.
> 2. ancestors of *t* (earlier members in the lineage) produced *F*.
> 3. there was selection for *T* because ancestors produced *F*.
> 4. where the selection for *T* because of *F* was in the recent past.

Adding this clause, the appendix, like the eyes of some cave-dwelling creatures, has no function. And feathers have more than one function.[20]

Etiological functions parallel functions from intentional design in allowing for change in function. Just as you can redesign a refrigerator, selection can redesign organs and other traits. A trait or part of an organism can even be redesigned so that it malfunctions according to the old design plan but functions properly according to the new design plan.

Both sources make sense of the function-accident distinction; both make sense of normal conditions and functioning normally; of the distinction between failure to operate normally and failure to fulfill a function; of the plurality of functions; of change in function and even the possibility of conflicting functions. One source comes from an intelligent agent sitting down and designing an item to solve a certain problem or produce a certain effect in a certain environment. Another comes from selection pressure acting on the features of a population over generations. I say two *sources* because of the extensive overlap between the two. I say two *kinds* because of the differences.

III. Plantinga'S Counter-Example

Two kinds or sources of function (purpose) and normal functioning (operating according to a design plan)? Design either with or without a designer? Plantinga will have none of this. For him there is only *one* kind and *one* source: intelligent, intentional design. This source is not only the ancestral home of the use of our teleological vocabulary, it is the *only* source of purpose and plans in the world; there are no purposes or plans from selective histories. And Plantinga has examples to prove it.

His central counter-example involves a Hitler-like madman, his henchman, and their persistent descendants over generations:

> A Hitler-like madman gains control:…he orders his scientists to induce a mutation into selected non-Aryan victims. Those born with this mutation can't see at all well (their visual field is a uniform shade of light green with little more than a **(p.78)** few shadowy shapes projected on it). When they open their eyes and use them, furthermore, the result is constant and severe pain, so severe that it is impossible for them to do anything except barely survive…. Their lives are poor, nasty, brutish, and short. [He begins] a systematic and large-scale program of weeding out the non-Aryan nonmutants before they reach reproductive maturity…. But then consider some nth generation mutant *m*. He is a member of a reproductively established family and has a certain reproductively established character (the relevant part of which involves his visual system). He has ancestors, and among his ancestors, there was a causal connection between that character and the way their visual systems performed, which accounts for the positive correlation of that character with that way of functioning among his ancestors…. [This] way of functioning conferred a survival advantage, [the madman and] his thugs and their successors were selectively eliminating those who do not display it. But wouldn't it be wrong to say that *m*'s visual system is functioning properly? Or that its function is to produce both pain and a visual field that is uniformly green? Or that the resistance medical technicians who desperately try to repair the damage are interfering with the proper function of the visual system?… it surely won't be the case that [his] visual system is functioning properly, or that its function is to produce pain and a visual field displaying nothing but a uniform expanse of green.[21]

The conditions for function and normal functioning on the etiological theory seem to be met, but surely the function of *m*'s visual system is *not* to produce a uniform expanse of green or cause severe pain; *m*'s visual system is surely *not* functioning normally. Counter-example!

Considered more abstractly, here is the example:

> *Step One*: Take an organ that is clearly adaptive, such as the human eye or the legs. Eyes are for seeing; legs are for locomotion. In normal conditions when functioning normally, eyes lead to keen vision, and legs help us move easily about our environment. Our eyes help us achieve a vast number of goals. Indeed, nearly half the human brain is devoted to vision. And through vision we find food, fight, flee, and find mates.
>
> *Step Two*: Take a population of humans. Then mutate one or more of the genes responsible for the organ from Step One in a subset of the population so that the organ ends up unable to perform its function, even in normal conditions. The mutated genes produce deformed, ineffective organs.
>
> *Step Three*: Select out the members of the population without the mutation. Stipulate as well that later descendants who lack the mutated gene continue to die off before reproduction so that the non-mutated genes fade away in the gene pool.
>
> **(p.79)** *Step Four*: Stipulate that the members with the mutated genes nevertheless survive long enough to have children, and their children with the inherited gene also continue to reproduce.
>
> *Step Five*: Note that the conditions for the mutated gene to count as an adaptation to the new selection pressures in the environment have been met. The etiological function of the altered visual system should now be the effect that explains why the organ is adaptive in the changed environment, and the organ should now be operating as it did during the selection process as its criterion for normal functioning.
>
> *Step Six*: Conclude that the result in Step Five is deeply counter-intuitive. The altered organ, intuitively, is horribly deformed and not functioning normally at all. And intuitively the altered organ has its previous function as its only function, and does not have the new effect that makes it adaptive as a function.

No naturalist can explain the correct application of the terms “function,” “purpose,” “design plan,” and “normal functioning.” Naturalism defeated!

Not so fast. First, one should clearly notice how artificial the example is. I don't mean the appeal to science fiction. Genetic engineering exists, and we might be able to write such mutations into our genetic code. And we all know the horrible experiments that Hitler and Hitler-like madmen have carried out on various populations. What makes the example so artificial is the idea that those with the mutation should go on reproducing. As Plantinga puts it, the victims' lives are “poor, nasty, brutish, and short.” Many will die well before reproductive age. Many will fail to reproduce if they live to that age. And those that do reproduce are not likely to be in a position to care for their children, who will once again have lives that are “poor, nasty, brutish, and short.” Many will choose not even to try to have children. The mutation is so deeply non-adaptive that the subpopulation, given our knowledge of how the world works and the adaptive advantage of normal human vision for human beings, is likely to diminish over time. Extinction of the subpopulation may even occur. And this is why Plantinga's stipulation that they will continue to reproduce and have descendants that inherit the gene is so fishy. Nature, left to its own devices, will reduce or even eliminate the subpopulation over time. Deeply maladaptive mutations don't last. Step Four seems bogus.

And this fishy stipulation is essential to the rhetorical effectiveness of the case. The visual system, so mutated, is clearly not “functioning properly.” The more maladaptive the altered trait happens to be, the stronger our intuition that the trait is not fulfilling its function and is not functioning (operating) as it should.

(p.80) So if maladaptive traits don't last, the altered visual system wouldn't meet the criteria for having an etiological function. Counter-example diffused.[22]

But even waiving the artificiality of the example, Plantinga seems not to have thought through the consequences for his *own* account of purposes and plans.

For consider the following question. Why did the madman and his henchmen mutate the non-Aryans in the first place? What reasons come to mind? Perhaps the pleasure that comes to the sadist through inflecting pain and suffering on others. The Hitler-like madman, his henchmen, and their descendants have altered a population to render their lives poor, nasty, brutish, and short, generation after generation. They have designed a gene to create an organ that causes suffering, and have done so for their own sadistic pleasures; they have redesigned the visual systems of non-Aryans for diabolic ends with a *non-epistemic* purpose and design plan.

But if that is so, the eyes and visual systems of the non-Aryans have acquired *new* designs with *new* purposes from intelligent design. They are designed so that when operating according to the new design a visual expanse of near uniform green will occur when the eyes are opened, accompanied by severe pain. Any non-Aryan with the redesigned visual system that does not operate that way is not operating according to the new plan. The redesigned visual systems are *functioning properly* according to the *new* design plan when causing green visual fields accompanied by severe pain.

So let us assume, for the moment, that the visual systems of the non-Aryans were originally intelligently designed with an epistemic purpose. Then when altered they are no longer capable of operating according to the *original* design plan. They clearly fail to function properly. But this is only so relative to the *old* design plan and purpose. Plantinga's case is simply one **(p.81)** where *redesign* leads to malfunctioning according to the *old* design, but proper functioning according to the *new* design.

To quote and paraphrase Plantinga, a “thing can obviously acquire a *new* purpose and *new* design plan”: I construct a device with a design plan for a purpose. Later I make a modification in the design plan for a different purpose. “It has acquired a new design plan. Relative to its old design plan, it is not functioning properly; relative to the new, it is.” Most drastically, I make major modifications for a contrary purpose. Then “when it functions properly according to the new design plan, it works badly according to the old.” Applied to the mutated non-Aryans, their visual systems “work in accord with the new design plan [but not with] the old design plan.”[23]

Plantinga is not entitled to assert that in *no* sense are the mutated visual systems functioning normally.[24] In constructing the example he guaranteed that the new visual system would have a new purpose and plan either from intelligent design or from artificial selection, or both. On his understanding of the source of purposes and plans, the new visual systems are, in a clear sense, functioning properly.[25] In Plantinga's example the jig is up at Step Two.

That said, can we construct a parallel example that doesn't involve intelligent design? It shouldn't be too hard: **(p.82)**

> The madman and his henchmen have taken over. They are systematically murdering non-Aryans. A radiation burst alters the DNA of a group of non-Aryans about to reproduce. The DNA they pass onto their children includes a mutated gene that substantially alters the building of their visual systems. Their visual systems are now as Plantinga has described. When the madman's henchmen approach these children, they let them live; they delight in the pain and suffering the children experience. The children are assisted by charitable Aryans or cruel henchmen, and the children start having children soon after reaching reproductive age. Their children acquire the immunity, though non-Aryans without the mutation are still hunted down and murdered. Over generations, the frequency of the gene that leads to the altered visual systems continues to increase in the subpopulation of non-Aryans. Those without the gene are hunted down and exterminated.

Here we have a case of variation, consequences, and heredity. So it looks like the conditions are met for the mutated trait to acquire a purpose and plan. But, intuitively, the altered visual system is not functioning properly. It's not working the way it is supposed to. It's clearly broken. If anything, it is malfunctioning. And it's obviously not fulfilling its purpose. The mutated non-Aryans can't see much of anything at all! Counter-example!

Again, not so fast. Just as items with purposes and plans from intentional design can acquire *new* purposes and plans that conflict with *old* purposes and plans, so, too, items with purposes and plans from selection can acquire new purposes and plans that conflict with the old. So there are two issues to address. First, is there a clear sense on the etiological theory of functions where the mutated visual system is not functioning properly, and not fulfilling its function? Can the etiological theory explain why it's obvious that in the example the mutated visual systems are not doing what they are supposed to do? Second, is it really counter-intuitive to imagine the altered visual system acquiring a new purpose and plan over time? If intelligent designers can assign new purposes and plans, why can't selection acting on variants in a population over time do the same?

To answer the first question, it is worth emphasizing that visual systems are a paradigm case of adaptations in nature. Even hard-core anti-adaptationists agree. On the etiological theory of functions, the human visual system is for accurate and reliable visual representation of various features of our environment. The etiological theory has no more problem explaining why the Aryans have normally functioning visual systems than it does explaining why the hearts of Aryans and non-Aryans alike have normally functioning hearts with pumping blood as a function: none.

So what does the etiological theory imply about the altered visual systems of first-generation non-Aryan children? According to the theory, their visual systems do not develop as they should. The mutated gene interacts with the **(p. 83)** other genes that build the visual system and causes the visual system to develop in the wrong way. As they grow and mature their visual systems are deeply malformed. Their visual systems are then not able to operate as they should. They do not function normally, and they don't fulfill their function. (A medical technician who desperately tried to repair their visual systems might get their visual systems to function properly; he might also get them killed.) According to the *old* purpose and plan, they are *malfunctioning*.

And since first-generation mutations have no new functions on the etiological theory, the first non-Aryan mutants, though they are now immune to the murderous intents of the madman and his henchmen, do not have altered visual systems with a new function or with a new standard for normal functioning. They simply have the old purpose and plan, a purpose they cannot fulfill and a plan they cannot follow. They haven't even acquired a new purpose or plan, unlike the one-off acquisition of new purposes and plans from intelligent design.

The etiological theory thus entails that the first-generation altered visual systems are not working properly and are not fulfilling their functions. For the first generation, the counter-example cannot even get off the ground. According to the old purpose and plan, their visual systems are not working properly.

So fast-forward a number of generations. Somehow or other the murderous henchmen have continually reproduced their kind and kept on murdering non-Aryans without altered visual systems; only altered non-Aryans have lived long enough to reproduce. Directional selection has produced evolution in the subpopulation. Given the murderous designs and overall effectiveness of the henchmen, stabilizing selection prevents new mutations that aid non-Aryans in getting around from finding a foothold in the gene pool. A visual system that doesn't cause a lot of pain when your eyes are open gets you killed rather quickly. A visual system that produces a nearly uniform expanse of green and hurts like hell keeps murderous henchmen at bay. Since the latter keeps you living long enough to reproduce, and the former only gets you killed, the altered visual system is not only more adaptive than the former, the genes that lead to the altered visual system have increased in frequency over time through selection. The altered visual system is thus an adaptation for immunity from the murderous designs of the henchmen. It keeps them away, and that explains why the altered system persists in the population. Selection has produced a new purpose and plan.

Given enough generations, the visual system of non-Aryans changes its purpose and plan. When the alteration first occurs, it is simply malfunctioning and does not fulfill its function. But over time it acquires a new purpose **(p.84)** and plan. Relative to the old purpose and plan, it is malfunctioning. Relative to the new purpose and plan, it functions properly. Redesign by natural selection has changed the purpose from keen vision to immunity from murderous henchmen.

Despite the artificiality of even this example, there are analogues in nature. First, there are cases where an organ loses its function due to a change in the environment. Cave-dwelling fish still have eyes, but don't use them for vision at all, and if exposed to light, they wouldn't be able to see a thing. The eyes are still there, but no longer have vision as a purpose. So just as the visual systems of some fish can lose vision as a function, so, too, the visual systems of some humans could lose vision as a function.

Second, there are cases of change in function and body plan that are at least accompanied by pain, even if the pain isn't instrumental to the selected effect. For example, walking on two limbs instead of four notoriously produces lower back pain. Adaptations result from a trade-off of costs and benefits. The benefits outweigh the costs, even where the costs involve a good deal of possible or actual pain. Pain itself is often functional; it leads us to avoid the cause of the pain.

Third, there are even cases of genes that produce immunities to diseases that are accompanied by other costs. For example, malaria can be quite a problem in various parts of the world, especially Africa. Many Africans have developed immunity to malaria. The gene that codes for the blood type is thus adaptive in their home environment. However, the gene also leads to sickle-cell anemia when a child inherits both recessive alleles from her parents. The gene doesn't cause the immunity by causing anemia; anemia is a side effect of the immunity. Causing anemia isn't what the genes are for. But even so, immunities often have a price to pay. Our made-up immunity to the henchmen is no different.

So in our new version of Plantinga's case without intentional design we have a parallel result to Plantinga's case with intentional design: an old purpose and plan is replaced by a conflicting purpose and plan. Relative to the old, the trait is not functioning normally or fulfilling its function. Relative to the new, it fulfills its purposes and functions properly. The counter-example is diffused.

IV. On Design and Designers

Plantinga's confidence in his example is driven, I think, by his confidence that the terms “purpose,” “plan,” and “proper functioning” analytically entail intelligent, rational, intentional design. If there were no powerful **(p.85)** intelligent designer and creator of living things, they would have no purposes, no plans, and nothing would count as proper functioning:

> [I]f naturalism were true [and there were no God]…neither human beings nor their component systems would function properly (or, for that matter, improperly). Fundamentally, that is because the notion of proper function really only applies to things that have been designed by conscious, purposeful intelligent agents.[26]

So if the systems and organs of humans have purposes and plans, they must have been designed by an intelligent, purposeful being:

> It looks initially as if the notion of proper function entails that of design, so that necessarily, anything that functions properly has a design plan. And it also looks as if necessarily, if a thing has a design plan, then either that thing or some ancestor of it was designed by a rational being.[27]

> Proper function requires design; but the only plausible designers for us human beings and our systems and organs would be God…[or] a high-ranking angel.[28]

We are now on very familiar ground. Plantinga's argument reduces to the old saw that things in nature are well designed; analytically the only way they could be designed is if a powerful intentional designer exists; so God exists as the intelligent designer and creator of all things.

If we give Plantinga the second premise, naturalists will balk at the first. Things may *seem* well designed, but that does not mean that they *are* (intentionally) designed; they may *seem* designed without *being* designed. Selective histories produce traits that *seem* well designed, that *seem* to have a designer, with a *seeming* purpose and *seeming* design plan, but in fact they are *not* designed, do *not* have a purpose or a design plan. It just looks *as if* they do. If Plantinga's second premise about the analytic connections between certain terms holds true, then his first premise is false. This point is all too familiar.

I prefer to reject the second premise and accept the first. I agree that the words “function,” “purpose,” and “design” are all ordinary English words with ordinary meanings. But I do not agree that it is analytic in ordinary English that an item has a function or purpose if and only if intentionally designed, or that an item can function properly if and only if it has a blueprint written up in advance by an intelligent designer. And for the very same reason I do not think it is analytic in ordinary English that an item has a purpose and plan if and only if selected for a certain effect over generations (see note 14). I see both intelligent design and selective history as **(p.86)** sources of purpose and plans, and not competing analyses of the meaning of teleological vocabulary in ordinary English.

These terms have thin senses readily grasped by competent speakers of English, correctly applied to items intelligently designed or selected for a certain effect. Just as ordinary folks use teleological vocabulary when talking about their kitchen appliances, so, too, evolutionary biologists use the same teleological terms when talking about the traits of living organisms. Neither misuse those terms; neither utter obvious analytic absurdities.

What Plantinga sees as analytic is really just a folk theory of how many items with functions come to have those functions. Intelligent designing is a source of purpose and plan with which we are all familiar, having designed devices to solve various problems ourselves. Absent an alternative theory for how an item can acquire a purpose and plan, it is only natural that we would apply our folk theory to any item we ran across that seems to have a purpose and plan. It was Charles Darwin's genius to uncover another source. The etiological theory of functions is thus a theoretical account of how items may acquire purposes and plans other than by intentional design. That it is not on the tip of ordinary English speakers' tongues when they use teleological vocabulary only shows that the theory is not widely understood. It doesn't show that it is false by virtue of the meaning of the words “function,” “purpose,” and “plan.”

Two other objections from Plantinga are now easy to dismiss. He thinks our systems and organs could have purposes and plans that have nothing to do with survival and reproduction. To quote and paraphrase Plantinga, “it is possible that a thing have a function (and function properly) even if that way of function confers no [survival enhancing propensity] upon its owner.”[29] He also thinks that an item needn't have ancestors either to have a purpose or plan. The “first telephone…could have had a proper function and could have been capable of proper functioning even if it had no ancestors at all.”[30] If there are two sources of purpose and plan, then of course both of these things are possible. God could assign functions to items that have no survival-enhancing propensity or items that lack ancestors. But for the very same reason selection could assign functions that have nothing to do with the intentions of some divine being. You don't need selection to assign functions if an intelligent being can do it; but you don't need an intelligent designer either when selection can do the work.

A compromise might be had between the first and second naturalist replies to the design argument: regiment teleological vocabulary when **(p.87)** describing the sources. When selection over generations does the work, use the terms “function,” “normal functioning,” and “normal conditions.” When intentional designing is the source, use the terms “intended purpose,” “design plan,” “proper functioning,” and “intended circumstances.”

The design argument could then be reformulated: the systems and organs of humans and other living organisms have intentional designs and design plans; so they must be created by a powerful intelligent designer, God. Naturalists would clearly reject this argument. The seemingly unobjectionable first premise is not so unobjectionable after all.

A thorough understanding of the two sources of purpose and plan diffuses Plantinga's counter-example and thereby undermines his version of the well-known design argument for the existence of God.[31]

References

Bibliography references:

Ariew, A., Cummins, R., and Perlman, M. (eds.) (2002), *Functions: New Essays in the Philosophy of Psychology and Biology* (New York: Oxford University Press).

Bigelow, J. and Pargetter, R. (1987), “Functions,” *Journal of Philosophy*. Reprinted in Buller (1999).

Buller, D. (1998), “Etiological Theories of Function: A Geographical Survey,” *Biology and Philosophy* 13: 505–27. Reprinted in Buller (1999).

Buller, D. (ed.) (1999), *Function, Selection and Design* (Albany, NY: SUNY Press).

Buller, D. (2002), “Function and Design Revisited,” in Ariew *et al.* (2002).

Godfrey-Smith, P. (1994), *“A Modern History Theory of Functions,” Noûs* 28: 344–62.

Griffiths, P. (2006), “Function, Homology, and Character Individuation,” *Philosophy of Science* 71: 1–25.

Levin, M. (1997), “Plantinga on Functions and the Theory of Evolution,” *Australasian Journal of Philosophy* 75: 83–98.

Lewens, T. (2004), *Organisms and Artifacts* (Cambridge, MA: The MIT Press).

McLaughlin, P. (2001), *What Functions Explain* (New York: Cambridge University Press).

Millikan, R. (1984), *Language, Thought, and Other Biological Categories* (Cambridge, MA: The MIT Press).

Millikan, R. (1993), “Propensities, Exaptations, and the Brain,” in her *White Queen Psychology* (Cambridge, MA: The MIT Press).

(p.88) Millikan, R. (2002), “Biofunctions: Two Paradigms,” in Ariew *et al.* (2002).

Neander, K. (1991), “The Teleological Notion of ‘Function,’ ” *The Australasian Journal of Philosophy* 69(4, December): 454–68.

Plantinga, A. (1993), *Warrant and Proper Function* (New York: Oxford University Press).

Plantinga, A. and Tooley, M. (2008), *Knowledge of God* (Malden: Blackwell Publishing).

Schwartz, P. (2002), “The Continuing Usefulness Account of Proper Function,” in Ariew *et al*. (2002).

Sober, E. (1984), *The Nature of Selection* (Cambridge, MA: The MIT Press).

Sober, E. (1993), *Philosophy of Biology* (Boulder, CO: Westview Press).

Notes:

(1) Plantinga gives another related argument against naturalism, the Evolutionary Argument Against Naturalism. I shall not state or discuss it here.

(2) Plantinga (1993: 196–7).

(3) Plantinga (1993: 214–15).

(4) Ibid. (194).

(5) Ibid. (13).

(6) Ibid. (21–2).

(7) Plantinga (1993: 26–7).

(8) Plantinga (1993: 26).

(9) Plantinga (1993) applies this point to an example from William Hasker. The example involves Geoffrey lacking vision due to a random genetic mutation, but acquiring the capacity to make his way around his environment via magneticolocation, a capacity only some marine organisms are known to have. Hasker claims that since Geoffrey's new ability is a mutation, it has no design plan. Not so fast, says Plantinga. He says Geoffrey may have acquired a new design plan by God's intervention, or the intervention of an intelligent agent:

> [Perhaps] Geoffrey has been given his new design plan by God or by someone else—some young and inept angel, perhaps—to whom God has delegated this task. True (as Hasker points out), you would not ordinarily expect God to revise Geoffrey's cognitive capacities in the direction of less rather than greater cognitive capacity, but of course that is by no means decisive. Perhaps there is some great good for Geoffrey that God can best achieve by making or permitting this revision. (30)

So just as a refrigerator can acquire a new design plan inconsistent with the old, so, too, we also can acquire a new design plan inconsistent with the old. Perhaps we are redesigned to make us better off or for some greater good. Or perhaps our redesign is for some malicious purpose by an evil being that makes us worse off.

(10) This example comes from the the UC Berkeley website on understanding and teaching evolution. For more on understanding evolution, take a look.

(11) The beetle example involves *directional* selection, where the frequency of a trait in a population changed over time due to selection pressures. *Stabilizing* or maintenance selection occurs when the frequency of a trait stays reliably constant over time. To continue our example, mutations that alter coloration are all likely to be less fit, and so apt to be selected out. A very dark green or very light green beetle may stand out and not survive long enough to reproduce. A yellow or red one is sure not to last long at all. Though directional selection produces changes in gene frequencies, and stabilizing selection keeps them reliably constant, the mechanism in both cases is the same: variation, heredity, and consequences.

(12) Being an adaptation is not the same as being adaptive. An adaptive trait is one that currently contributes to fitness. Adaptations evolve, in part, by being adaptive, but also by being better adaptive than their variants. So adaptations must have had variants, but adaptive traits, as such, need not. In addition, an adaptive trait might arise by a random mutation out of the blue, and so an adaptive trait need not even result from a historical process. And though being historically adaptive explains how adaptations evolve by natural selection, an adaptation might not be currently adaptive, for the circumstances may have changed so that the trait no longer contributes to fitness. Sober (1984: 100) and Sober (1993: 83–4).

(13) There is a lively debate among naturalists over the best way to understand nature as a source of function and design. The present debate is thus between Plantinga's view and the view I favor. David Buller (1998, 2002) has proposed a weakening of the etiological account so that selection for the trait is not required but only that the trait contribute to the fitness of the organism or larger containing system as the organism or system undergoes selection. History is still required, but variation on the functional item is not. John Bigelow and Robert Pargetter (1987) have proposed a non-historical account that ascribes functions to traits that have a propensity to be selected for. So variation is still required, but history is dropped. Tim Lewens (2004) has proposed that functions are just adaptive effects, regardless of selection. So Lewens drops both history and selection. Peter McLaughlin (2001) has suggested that functions are not tied to selection and reproduction, but self-maintenance of the individual organism. For Buller a functional item is a trait that is there by virtue of being adaptive in the past. For Bigelow and Pargetter a functional item is a trait that is currently adaptive and undergoing selection. For Lewens a functional trait is a currently adaptive trait that need not be undergoing selection. For McLaughlin functions are not tied to fitness. For discussion of the possible compatibility of historical accounts of functions and present contributions to fitness or maintenance accounts, see Millikan (1993, 2002) and Griffiths (2006).

(14) Advocates of the etiological theory include Peter Godfrey-Smith, Ruth Millikan, and Karen Neander, among others. Larry Wright's 1973 article moved accounts of functions in the explicitly historical direction, but his view was not explicitly historical. For a collection of articles, see Buller (1999).

Does the etiological theory give the meaning of the word “function” in ordinary English, or its meaning in the mouths of professional biologists? No. Though biologists use “function” all the time, especially in more popular writings, “function” is not a technical, theoretical, or explanatory term in evolutionary science. As a result biologists tend to use the word “function” freely in all three of these senses, sometimes for purpose, sometimes for capacity, and sometimes for how a gene or trait operates, or how a trait contributes to fitness. They use the term roughly as it is used in ordinary English. They simply see, when it comes to the first and third uses, another source of purpose and plans: natural selection.

Though some advocates of etiological accounts of function have seen their accounts as giving the meaning of a use of the word “function,” the etiological account as I understand it is not an analysis of the meaning of the word or of the concept associated with that use of the word. Rather, the etiological account of function is a theoretical account of one source of purposes and plans.

(15) Millikan (1984).

(16) “Normal” in this context does not mean “usual” or “average.” The way an item normally functions may not be the way that it works on average. Normal conditions are not necessarily the conditions where it usually is. Millikan (1984: 33–4).

(17) There is an interesting difference here between functions from intentional design and functions from selective histories. An item with an etiological function will, *ceteris paribus*, fulfill its function when functioning normally in normal conditions. It's guaranteed that design plans for items with etiological functions are good. But there is no guarantee that design plans for items with intended functions will be good. The designer may have been ignorant of what was required to come up with a good design. The item, once designed, may fail to fulfill its purpose, even when functioning according to the design plan. Just think of the dustbin of failed inventions. For failed inventions, there is no explanatory connection between its function and fulfillment of its function in past circumstances, for it does not have its function because of past success. Or an item with an intentionally designed plan may fulfill its intended purpose, but only do so accidentally. It worked on this occasion, but only by lucky accident. Design plans, consciously assigned purposes, and normal conditions bear no deep explanatory individuative relations to one another. For items consciously designed, these notions are atomistically distinct.

(18) This answers Plantinga's objection that naturalist theories are circular. Plantinga (1993) argues that Bigelow and Pargetter's propensity account of functions is circular in helping itself to an account of natural habitat (normal conditions). He then generalizes to all naturalist accounts. Plantinga is on solid grounds when criticizing Bigelow and Pargetter on this score. Neander (1991) and Millikan have made the same point. But Plantinga is wrong to generalize to the etiological theory.

(19) Godfrey-Smith (1994).

(20) Cf. Millikan (1993, 2002); see also Schwartz (2002) for further refinement.

(21) Plantinga (1993: 203–4 and 206).

(22) I suppose we could imagine a way out for Plantinga. To keep the mutation in the gene pool over generations, the madman and his henchmen have descendants that keep up the malevolent murder of descendants who lack the altered visual system. How do they keep those with the altered visual system reproducing? I suppose they build camps for the victims to reproduce. They provide food and shelter and medical treatment. If the pain gets too severe for the non-Aryans to remain conscious, they may even provide anesthetics to keep the non-Aryans alive and reproducing. The non-Aryans live lives that are poor, nasty, brutish, and short, but long enough to reproduce, generation after generation, as the Aryans keep the population stable over generations.

In his reply to Michael Levin (1997: 91) on this point, Plantinga (2008: 28–9) imagines the generosity of other Aryans. They keep the mutants safe and alive. They provide shelter. They assist in their reproduction. They even mate with some of the non-Aryans, producing mutated offspring. So maybe the henchmen build concentration camps, or charitable Aryans build nursing homes. Either way, somehow or other the non-Aryans with intuitively deeply maladaptive visual systems keep on trucking.

(23) Plantinga (1993: 26). See note 2.

(24) Michael Tooley independently (Plantinga and Tooley 2008: 186) notices this point, but does not pursue it. On the other hand, Tooley does effectively dispatch Plantinga's contention that health and proper functioning are necessarily coextensive, thus showing that intuitions about whether an individual is healthy can't drive an account of whether the individual's systems and organs are functioning properly. Tooley imagines a virus entering John's body that destroys a mechanism in the cells. Then the cells are no longer functioning properly. However, this mechanism was responsible for aging. Perhaps God or Mother Nature designed us to age and die. Or perhaps the change blocks the pain that accompanies terminal illness. Those dying of a terminal illness can thus seem fit as a fiddle. After the infection, John is much healthier than before, but his cells do not operate according to their design plan.

(25) Since intelligent design is doing all the work through genetic engineering, keeping the population reproducing doesn't matter. One scientist engineering one visual system is enough to confer on the visual system a new purpose and design plan. Plantinga included the murderous weeding out of non-mutant Aryans over generations in order to create a counter-example to the selective history account.

Michael Levin argued that the example produced a new purpose and plan through artificial selection. Levin does not discuss the point about genetic engineering. Plantinga's reply (2008: 28–9) to Levin is to note that the mutation can get out of control and spread through the entire population, presumably by virtue of Aryans mating with the mutated non-Aryans, and somehow overcoming all of the murderous Aryans (how is that supposed to happen?). How this virus-like spreading is supposed to answer Levin's argument escapes me. Did the selection not occur? And if not, then Plantinga's case is just a case of mysterious genetic drift, and not a counter-example.

(26) Plantinga (2008: 20).

(27) Plantinga (1993: 198).

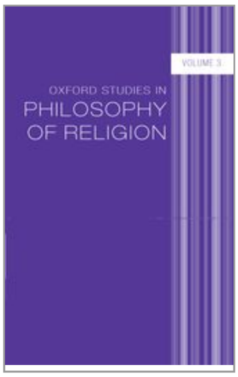
(28) Plantinga (2008: 20).

---

(29) Plantinga (1993: 209).

(30) Ibid. (203).

(31) I am grateful to the referees for Oxford Studies for their comments on the penultimate draft.



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

A Puzzle about Hypocrisy

Frances Howard-Snyder

DOI:10.1093/acprof:oso/9780199603213.003.0006

Abstract and Keywords

This chapter defines hypocrisy as moral inconsistency where the subject does x, while making some verbal or behavioral expression of the claim that x is wrong. It considers purported counterexamples and argues that the original definition survives most of these, although it may need a little tinkering. The second half deals with the concern that moral realism has a difficult time explaining why hypocrisy is so objectionable, while constructivist moral theories can more easily explain this. In response the author argues that, even if realism is true, the hypocrite is guilty of the special vice of knowingly doing what is wrong. Either she is acting wrongly or she is wrongly condemning or blaming others or placing unnecessary burdens on them. This is true even (or maybe especially) in cases where it isn't clear what the right action is. If we think of hypocrisy as a vice (rather than a wrong action) this seems to illuminate its special sort of badness.

Keywords: hypocrisy, moral realism, vice, inconsistency, constructivism

I. Examples of Hypocrisy and Charges of Hypocrisy

In our age and in ages past, hypocrisy has been regarded as worthy of a particular sort of contempt and revulsion. Jesus reserves his strongest criticisms for hypocrites, suggesting that even the gravest sins of other sorts can be forgiven but not hypocrisy:

> But woe to you, teachers of the law and Pharisees, you hypocrites! You shut the kingdom of heaven in men's faces: You yourselves do not enter, nor will you let others enter who are trying to. (Matthew 23:13)[1]

> Woe to you, teachers of the law and Pharisees, you hypocrites! You give a tenth of your spices—mint, dill and cumin. But you have neglected the more important matters of the law—justice, mercy and faithfulness. (Matthew 23:23)
>
> You hypocrites! Isaiah was right when he prophesied about you:
>
> > “This people honor me with their lips
> > but their hearts are far from me.” (Matthew 15:7, 8)

Examples in literature abound. William Shakespeare's Malvolio (*Twelfth Night*), Polonius (*Hamlet*), and Angelo (*Measure for Measure*) are all guilty of hypocrisy. *The Merchant of Venice* includes dramatic examples. Shylock refuses to grant mercy to Antonio and then begs for mercy for himself. Some critics of the play have seen Antonio, Bassanio, and Portia as hypocrites too. They plead for mercy for Antonio on the grounds that mercy is divine and “like the gentle rain from Heaven” and then treat Shylock unmercifully. Uriah Heap (in Charles Dickens's *David Copperfield*), Mr. Collins (Jane Austen's *Pride and Prejudice*), and **(p.90)** Tartuffe (Molière's *Tartuffe, or the Imposter*) are amongst a host of literary hypocrites.

Hypocrisy is not just an old-fashioned vice. Accusations of hypocrisy are common in our culture. We frequently hear news reports of politicians or preachers who have been outspoken critics of gay marriage, who are themselves outed by gay prostitutes. There's even a website called “Where's the Hypocrisy?” where people write in with examples. There's a section for Republicans: Al Gore has a big house that consumes a lot of energy; and for Democrats. Here's Gail Collins writing about Mitt Romney, the Republican former Governor of Massachusetts:

> Now Mitt Romney—there's a man who can put the pep into pander. “I am pro-family on every level, from personal to political,” he told the summiteers…. He reeled off anti-abortion pledges—not just the requisite anti-Roe Supreme Court nominees, but promises to “oppose abortion in military clinics, oppose funding abortion in international aid programs and I will work to ban embryonic cloning.” He was almost as impassioned as he was during his Senate race against Ted Kennedy when he talked about the “dear, close family relative who was very close to me” who died from an illegal abortion and his firm conviction that “we will not force our beliefs on others on that matter. And you will not see me wavering on that.”[2]

The November 2007 edition of the *Atlantic Monthly* had a scathing discussion of Hillary Clinton by Caitlin Flanagan, who describes a book Hillary wrote about her cat Socks, in which she tells readers to follow her example in caring for their pets, in not taking the responsibility of a pet lightly, etcetera. We later learn that Socks was given up for adoption soon after the Clintons left the White House.[3]

II. Four Types of Hypocrisy and Suggestions about What They Have in Common

So what is hypocrisy and why is it so objectionable? In “Hypocrisy and Moral Seriousness,”[4] Roger Crisp and Christopher Cowton identify four different kinds of hypocrisy which they call pretence, blame, inconsistency, and complacency, not because all instances of blame and so on are examples of hypocrisy but just that there are blaming, pretence, etcetera, types of hypocrisy.

(p.91) The kind of pretence Crisp and Cowton have in mind is a matter of pretending to virtue (or someone's idea of virtue). Here they cite Molière's Tartuffe as the paradigm example of a hypocrite. This is unlike some other kinds of hypocrisy in that this sort of hypocrite need not be condemned by his own (believed) standards.

The kind of blame the authors have in mind is, in its simplest case, when someone blames someone else for a fault he himself possesses. For example, Senator Joe McCarthy—it is said—publicly castigated gays for immorality although he was himself homosexual. When Jesus asks whoever is without sin to cast the first stone, he is reminding his audience that they are condemning a woman for a (kind of) sin that they themselves possess. But it isn't always the case that the blamer possesses the (very) same fault. In the Sermon on the Mount, Jesus asks “How can you say to your brother, Let me take the speck out of your eye, ’ when all the time there is that plank in your own? You hypocrite! First take the plank out of your own eye, and then you will see clearly to take the speck out of your brother's” (Matthew 7:4–5). So, blaming someone for doing X when one does X oneself, or blaming someone for doing X when one does Y oneself and X is wrong only if Y is wrong in a similar way or is worse than X, both count as this sort of hypocrisy.

The sort of inconsistency Crisp and Cowton have in mind is simply a matter of claiming that a certain sort of behavior I say is wrong, yet I do it myself. In Matthew 23:13, Christ advises his listeners to pay attention to what the Pharisees say because they are in the seat of Moses, and goes on, “But do not follow their practice: for they say one thing and do another.” This may not involve either blaming (if no one else is yet guilty of the proscribed activity) or pretence. In some cases, the agent is upfront about the inconsistency. Imagine someone very powerful who doesn't need to ingratiate himself, or a parent who admits to some vice but encourages his children to avoid it. Nicotine addicts might preface their instructions to their children not to smoke with, “I know I'm being hypocritical but…”

As an example of the fourth sort of hypocrisy, the authors cite the Pharisees again (at least as described by Jesus in the New Testament): “Alas, for you, lawyers and Pharisees, hypocrites! You pay tithes of mint and dill and cumin: but you have overlooked the weightier demands of the Law, justice, mercy, and good faith. It is these you should have practiced, without neglecting the others. Blind guides! You strain off a midge, yet gulp down a camel.”

These Pharisees count as hypocrites because they take morality seriously in very unimportant ways, ignoring its demands where their fulfillment appears costly. Apparently the stereotypical Victorians were like this, too, or at least that subset of the Victorians who cared about trivial matters like not **(p.92)** exposing a woman's ankle in public (and even sometimes covered up the “legs” of pianos) but did not care about the morally serious matters like the suffering of children in workhouses or abused young girls in brothels. Crisp and Cowton argue that these Pharisees may not be guilty of pretense or inconsistency or blame. Their concern for morality is hypocritical because it is extremely undemanding and they are unwilling to reflect on it.

So, what do these four kinds of hypocrisy have in common? The first three seem to have a common element: a kind of moral inconsistency where the subject, on the one hand, makes some verbal or behavioral expression of the claim that X is wrong while, on the other hand, doing X. It seems important that the expression of the claim that X is wrong must have an audience. Simply believing that X is wrong and doing X is not a clear case of hypocrisy.[5] Dan Turner writes:

> All the examples we have considered (and all others I am aware of) have at least the feature of conflict or disparity, and this feature seems to be universally accepted as essential. Every disparity or conflict requires opposing sides: thus words vs deeds, pretended beliefs vs genuine beliefs, beliefs vs desires; all seem appropriate forms in terms of which to generate hypocrisy.[6]

Crisp and Cowton find the fourth category of complacency difficult to fit into this framework. But I think they are mistaken about this. The paradigm example they offer of such complacency is the Pharisees who make a great fuss about exact obedience to the Law in trivial details like the tithing of mint and cumin, but ignore morally important matters such as compassion and forgiveness. I believe that, insofar as this makes them guilty of hypocrisy, it *is* because they are inconsistent. Their concern about the tithing of mint and cumin is derived from and supported by the Law of Moses, which they appeal to as their absolute authority. But they are not living by the important tenets of that Law. They claim to be following God, but are ignoring the considerations that God cares about most. Something similar might be said about the Victorians or, at least, those Victorians described two paragraphs earlier. They presumably appealed to Christianity to justify their concern about propriety; and Christianity, of course, has more to say about widows and orphans and the poor generally than about ankles and knees. The contemporary bumper stickers "Who Would Jesus Bomb?" and "Who Would Jesus Torture?" express similar charges of hypocrisy against a **(p.93)** certain sort of contemporary Christian who invokes Biblical authority on some matters but ignores it in other, more important matters.

Consider by contrast, a non-religious society that is set up to monitor correct usage of the English language. Its members write angry letters to *The New York Times* about split infinitives and dangling modifiers, but do not express any outrage at the oppression in Myanmar or the genocide in Darfur. These folk might be guilty of triviality, but I don't think that we would judge them to be hypocrites.

So, to conclude so far: all four kinds of hypocrisy that Crisp and Cowton identity seem to share an inconsistency between one's avowed or apparent standards, on the one hand, and one's behavior, on the other—where one's standards condemn the very behavior that one practices.

III. Apparent Counter-Examples

A worry: this sort of inconsistency between one's standards, or avowed or apparent standards, and one's behavior might not be sufficient for hypocrisy. Here are some kinds of behavior that critics often want to distinguish from hypocrisy:[7]

1. *Weakness of will*: If I say, "Don't smoke" and I believe that smoking is wrong, and I genuinely want to stop smoking, but I find myself still smoking, my behavior is inconsistent with my expressed moral norms, and yet, we typically want to classify this as weakness of will and not as hypocrisy.

2. *Change of mind*: If I argue that a certain kind of act is wrong at one time, change my mind at a later time, and then am observed to be doing this same kind of act, my later behavior is inconsistent with my earlier expressed norms, but this is usually not hypocrisy.
3. *Morally appropriate distinctions*: Parents sometimes tell their children not to do certain things, and then they (the parents) engage in these activities themselves. The combination of instruction not to do a certain activity A with the action of doing A looks like the kind of inconsistency that exemplifies hypocrisy. Yet these parents often defend themselves by pointing out salient differences between themselves and their children. Sometimes these distinctions do seem to exonerate the parent from the charge of hypocrisy. For **(p.94)** example, if I tell my nine-year-old children not to drive a car or wield a carving knife, but I go ahead and do these things myself, I am not being hypocritical. There are important differences between children and adults that make driving a car at nine wrong but driving a car at some age greater than sixteen and with a license not wrong.
4. *Saints in disguise*; People who appear to adopt a certain set of moral principles in order to undermine a wicked form of tyranny may exhibit inconsistency but do not seem guilty of hypocrisy. If Oskar Schindler (as depicted in the movie *Schindler's List*) acts as if he accepts Nazi morality in order to hide his true agenda of rescuing Jews from concentration camps, we don't judge that he was guilty of hypocrisy, even though his actions contradicted the morality he espoused to the Nazis around him.
5. *Retaliation or punishment*: If society or its agents kill a killer, they are sometimes accused of inconsistency or hypocrisy. But this seems overly simplistic. We don't make similar complaints when society deprives a kidnapper of his freedom or a thief of his wealth. Justice often involves doing something to the wrongdoer that is exactly like or similar to the wrong act. This is true on the interpersonal level as well as on the societal level.

Suppose I ask for a few days grace to pay back a debt that I am having trouble repaying. You refuse and insist that I repay the money on time. Sometime later our situations are reversed and you are indebted to me. You ask for an extension. Remembering our earlier interaction, I refuse. Now, there is a kind of inconsistency on my part here: I earlier (let's say) argued that it would be morally right for you to give me an extension. And now I am refusing to give an extension. But it seems that I am not exhibiting any vicious hypocrisy.

> 6. *Pharisees in reverse*: Imagine someone who appeals to the Law of Moses to justify urging others to love their neighbors as themselves, and to love God with their whole hearts and minds, and who acts on these himself, but who ignores the minor details of the law—such the tithing of mint and cumin. This person would be morally inconsistent but we wouldn't judge him to be a hypocrite.

Let's take these apparent counter-examples in order. In the first case of weakness of will, we have a simple case of inconsistency between the agent's avowed standards, on the one hand, and her practice, on the other. Is this hypocrisy; **(p.95)** and if not, why not? Here's a simple suggestion.[8] Each case of weakness of will is either a case where the agent *can* do otherwise or a case where she *cannot* do otherwise.[9] Insofar as the agent cannot do otherwise, the principle that “ought” implies “can” gets her off the hook. She hasn't acted wrongly. But perhaps what's wrong is the act of criticizing or blaming or punishing others for behavior she herself exhibits. She *can* avoid that, after all. But this does not seem wrong in every case. Think of a drug addict urging children to stay off drugs. Even though she can't help backsliding now and then, she seems to be acting responsibly in encouraging others to avoid her mistakes. One way to avoid the charge of inconsistency here would be to say that the real principle she is avowing is, “Do not do drugs if you can avoid it!” In that case, there is no inconsistency in her discouraging children from doing drugs (supposing they haven't yet become addicted) and her doing drugs herself.

At least some cases of weakness of will are cases where the agent *can* do otherwise. (“Like you, I really ought to give to famine relief, but that Armani sweater was just so darn adorable!!”) In these cases, perhaps we do have cases of hypocrisy. Presumably, one doesn't get off the charge of hypocrisy by pointing out that one was faced with resistible temptation. Now, there may be borderline cases, where the temptation is very powerful, not absolutely irresistible but very, very hard to resist. In this sort of case, we may say that the agent is guilty of hypocrisy but it is forgivable because of the difficulty of avoiding it. How much leeway we are prepared to give the agent in such a case will depend on how far we think difficulty excuses an agent. Whether the agent counts as a hypocrite may also depend on how generous and forgiving she is of others who are faced with similarly difficult temptations.[10]

Consider the second case of someone who changes her moral principles over time and acts accordingly.[11] At least some such cases don't seem to **(p.96)** count as hypocrisy. Consider the young person who believes that white lies are wrong, avoids them, and criticizes others who utter them. As she grows older she becomes more sensitive to the hurt that blunt truth can cause, changes her views about the ethics of white lies, and occasionally utters such lies herself. This is not hypocrisy. But it may not be a case of inconsistency either. At no point in the story does our agent say that white lies are wrong and practice them. In her youth, she said that white lies were wrong and didn't tell them; in middle age, she says that white lies are sometimes permissible and sometimes tells them. This is no more inconsistent than someone who first believes that Los Angeles is the capital of California and later comes to learn that it is not.

At the same time, some changes of moral opinion do look like examples of hypocrisy: the politician who at one time is pro-life (when he is running for election as the Republican nominee for president) but who earlier defended abortion rights (when he was running for election in a liberal state). Or consider a young woman who is in favor of preferential hiring in graduate school, but then becomes opposed to it later in life when she is an employer and the mother of (white) sons. These sorts of changes are often described as hypocrisy. Perhaps the thought behind the charge is that the change is so convenient that the agent must be guilty of either insincerity or self-deception at one or both times of her life. So, this case again fits with the “inconsistency hypothesis.” If an agent's attitudes genuinely and honestly change over time, she is not inconsistent. If she is lying to others or herself about her change of attitude, then she is guilty of some sort of inconsistency, and as such, is guilty of hypocrisy. In “convenient” cases, as described above, the expectation is that she is probably guilty of insincerity or self-deception.

Turning to the third kind of apparent counter-example, something similar can be said about offering distinctions. Sometimes they look like hypocrisy and sometimes they don't. Suppose that you and I are playing chess. I make a silly move, and ask you to let me take it back. You graciously agree. A little later in the game, you hang a queen, slap your forehead and ask if you can take it back. I say, “No. Touch move.” This seems like a paradigm case of hypocrisy. But if you call me on it, I might point out that you're a chess master, and I'm a chess novice, and as such, different rules of chess etiquette apply to us. This is (maybe) a reasonable distinction to make. Similarly, parents who drive but tell their children not to drive can point out salient differences between themselves and their children—better judgment, coordination, more developed brains, compliance with the law, etcetera.

By contrast, consider a man who is outraged when he discovers that his wife had sex before their marriage (such as Angel Clare in Thomas Hardy's **(p.97)** *Tess of the D'Urbervilles*). When she points out that he did the same thing, he argues that the cases are very different since he is a man. Offering this distinction doesn't get him off the hypocrisy hook. Parents who curse and then tell their children not to curse and justify their apparent inconsistency by pointing out that certain choice phrases are “adult words” seem to occupy a curious borderline between the hypocrite and the non-hypocrite. This kind of distinction-mongering can avoid inconsistency. The person accused of hypocrisy can say: “The rule I was invoking was, ‘Don't A in B circumstances.’ I did A. But it wasn't in B circumstances. You on the other hand are A-ing in B circumstances.”

There is no straightforward inconsistency between the action prohibited by the agent's expressed rule and the behavior the agent exhibits. However, an objector can argue that B-circumstances do not make a moral difference; and that whatever makes A wrong in B-circumstances makes it wrong in most non-B circumstances too.

Which distinctions will enable the speaker to avoid hypocrisy and which don't? One thought is that only genuinely morally correct differences will work. It is only because it is true that chess novices deserve a break and chess masters don't that it was okay for me to ask for a take-back and then to refuse to give one; whereas, fornication by men is no worse than fornication by women. It seems that one might be able to avoid being a hypocrite if one made an honest mistake. How to spell this out is tricky. Simply believing that some distinction makes a moral difference doesn't seem enough. People can have horribly morally pernicious beliefs. On the other hand, getting it absolutely right doesn't seem necessary either. A third intermediate possibility is that the distinctions have to be plausible candidates for morally relevant distinctions—to allow for honest, reasonable mistakes. This rather vague notion would allow for honest, reasonable mistakes, would not allow for outrageous examples of bad faith, and would leave the borderline around hypocrisy rather vague.

What about the Oskar Schindler saint-in-disguise character in the fourth counter-example? Perhaps the simple fact of the matter here is that hypocrisy is by definition bad. If some behavior is just like hypocrisy but is very heroic or noble, then we refuse to call it hypocrisy because a necessary condition on something's being hypocrisy is that it is bad. This may explain our verbal tendencies. It is not a very deep answer, I'm afraid.

How about retaliatory inconsistency, as in the fifth counter-example? Perhaps this is a case where there is a legitimate or possibly legitimate distinction. For example, people often charge advocates of capital punishment with something like hypocrisy: “You say killing is so bad that you're going to kill someone for it. Don't you know that two wrongs don't make a right?” **(p.98)** In response, it seems appropriate to point out that there are important differences between killing a killer and killing an innocent person (or killing a person who has been found guilty of murder by due process and one who has not). Similarly, we wouldn't accuse someone of hypocrisy who killed in self-defense on the grounds that she did exactly what her victim was trying to do to her. In more ordinary, day-to-day cases, we find people treating one another in these ways all the time. You refused to help me when I needed help. So, I refuse to help you. You broke promises to me repeatedly. So, I feel okay about letting you down today.

This attitude is expressed in the parable of the unmerciful servant (Matthew 18:23–35). In this story, Jesus describes a servant who asks for and receives forgiveness of a large debt from his master; the same servant then turns around and refuses to forgive a smaller debt from a fellow servant. The master is furious with the servant and punishes him to the full extent of the law.

Immanuel Kant hints at a similar idea[12] when he suggests that a retributive account of punishment is a natural extension of the categorical imperative. By stealing, an agent is in effect allowing others to steal from him; by killing, an agent is in effect allowing others to kill him. Shakespeare invokes this idea in a number of places, including *Henry V*, Act II, scene 2, where Henry converses with three of his subordinates. He tells them that he has just pardoned a man who had criticized him while drunk. The three, Cambridge, Scroop, and Grey, all object that this is too merciful. Henry then confronts them with the evidence of their treason. When they beg for mercy, he responds like this:

> The mercy that was quick in us but late,
> By your own counsel is suppress'd and kill'd:
> You must not dare, for shame, to talk of mercy;
> For your own reasons turn into your bosoms,
> As dogs upon their masters, worrying you.

In *Hamlet*, we get the same idea. When Polonius tells Hamlet that he will use the players “according to their desert,” Hamlet responds:

> God's bodykins, man, much better: use every man
> after his desert, and who should 'scape whipping?
> Use them after your own honour and dignity: the less
> they deserve, the more merit is in your bounty.
> Take them in.

(p.99) If we treat the players according to their desert, that is how we will be treated, and that will be bad for us.

And, of course, in *The Merchant of Venice*, this idea becomes the theme of Act IV, scene 1. Shylock has the legal right to take a pound of flesh from Antonio (given the contract they earlier made.) Portia, disguised as the lawyer, Balthazar, repeatedly asks Shylock to show mercy. Shylock repeatedly refuses, citing the importance of following the law. At the last moment, Portia turns the tables on Shylock showing that he himself is in violation of the law. When he asks for mercy, she repeats his words to him, echoing the parable from Matthew, with its central claim that we all need forgiveness and mercy, and those who refuse to give it will not receive it.

So, at least in some such "retaliatory cases," it seems appropriate to defend oneself against the charge of hypocrisy by pointing out that the punishment or retaliation is importantly different from the original offense.

The sixth counter-example of "Pharisees in reverse" who obey the important central difficult parts of the law, but ignore the trivial parts, create some difficulty for the view that hypocrisy is simply a matter of inconsistency between one's avowed morality and one's behavior. But if we note, as I have earlier, that agents seem to be clearer cases of hypocrisy when they are placing burdens on others—by blaming or punishing or threatening with Hell, etcetera, than when they are simply encouraging others in a caring compassionate way to do better—then perhaps we can add an extra condition to our account of hypocrisy. Hypocrisy is an inconsistency between avowed morality and behavior (or perhaps between two different pieces of avowed morality) where the intent or effect is to pressure others to take on the burdens associated with morality while failing to take on such burdens oneself. The Pharisees in reverse and the saints in disguise are taking on very large burdens to further their moral goals, probably more than most of those around them. So, they do not count as hypocrites.

IV. A Puzzle and an Argument about Hypocrisy

Even if we grant that hypocrisy is by definition bad, this doesn't answer the question of *why* it is bad. Supposing the term "hypocrisy" to be a thick concept, that has evaluative and non-evaluative components, we still need to address the question of why the non-evaluative component is considered almost always to be so very bad.

(p.100) The data suggest that hypocrisy is universally condemned. As I mentioned above, whatever their view of homosexual sex, people tend to regard as objectionably hypocritical someone who publicly condemns homosexuals but privately engages in homosexual activity; similarly, whatever their view about conservation, people tend to regard as objectionably hypocritical someone who publicly condemns polluters, but privately profits from a highly environmentally hazardous business.

This suggests that we should expect proponents of different moral theories to agree that it is seriously wrong.

Here's a puzzle. Different theories do not agree on the badness of hypocrisy. Suppose that moral realism is true, and that there are moral facts that are not made true by our moral choices (although it is consistent with this position that they be made true by God's choices). Of course, these moral facts may in part be made true by facts about our needs and desires and by what hurts us, but let's stipulate that the moral facts are not made true by our individual or collective moral choices. Let's call this theory (or collection of theories) M. Utilitarianism, and most other forms of consequentialism would count as instances of M, as would Rossianism.

Now suppose that we have an agent engaging in a classic example of hypocrisy: Harriet says to John, “You ought to do A,” and maybe even scolds him for failing to do A. But she herself does not do A. Let's contrast Harriet with Moira (the moralist) who says, “You ought to do A” and does A herself; and with Lindsay (the morally lax person) who does *not* say, “You ought to do A” and does not do A.

Intuitively, we want Harriet to turn out worse than Moira and worse than Lindsay (*ceteris paribus*).

You can imaginatively fill in “A” as you like, but I want to consider cases of different A's with different moral statuses.

Now, consider four possibilities:

1. A is morally obligatory.
2. A is wrong.
3. A is not morally obligatory but A is supererogatory.
4. A is neither obligatory nor supererogatory. It's neither better nor worse if people do A. It's a matter of indifference.

Let's first consider case (1) where A is morally obligatory (in fact and according to M—since M is true). In this case, what Harriet says is true: John really ought to do A. If we suppose that there are no morally significant differences between John and Harriet in any of these cases, then Harriet acts wrongly in failing to do A. Since M is the true moral view, it says that Harriet has behaved wrongly in failing to do A. So, M can condemn Harriet. **(p.101)** The problem is that Harriet's combined utterance and action are better than the combined behavior of Lindsay, because if it's good that people do A, then it is better that Harriet get John to do it than that she does not.[13] (It may happen, as a matter of contingent fact, that if John finds out that Harriet is a hypocrite, he won't do it. But let's suppose he does not find out.) So, M doesn't rank hypocrisy as especially bad.

On (2) what Harriet says is false, and it's bad that she is conveying moral falsehoods to people and especially if she succeeds in getting John to do A. So M can condemn her. The problem here is that Harriet turns out to be better from the point of view of M than Moira, because A is better off not done. So, in this case, again, M doesn't properly rank hypocrisy.

On (3) for example, suppose that giving a significant amount of your money to charity is great but not morally required. In that case, it is better that Harriet gets John to do so than not. It would be better still if she gave the same amount herself, but it is not obligatory. Once again, M judges Harriet's behavior to be better than Lindsay's. So, in this case, again, M doesn't appropriately condemn hypocrisy. I think M would judge that Harriet's behavior was imperfect, in attempting to create a belief in John that was in fact false (in this, Moira would be in trouble, too), but on the other hand, the effect of getting John to do A would be good, so the theory would judge that her behavior was in that respect better than Lindsay's.

On (4) suppose that A is wearing a hat in church. It doesn't hurt if you do it, and it doesn't hurt if you don't. It seems that Harriet is unnecessarily making John uncomfortable. But it turns out that Moira is unnecessarily making John uncomfortable, too. It is a minor wrong but it is one that Moira and Harriet share.

So, on each of our four suppositions about A, Harriet is ranked ahead of one of the more consistent women in our story. M, you recall, named any of a group of theories that were moral realist, where the rightness or wrongness of acts did not depend in any way on individual or collective judgments of agents.[14]

(p.102) Let's now consider some other theories. What unites these is that they are in various ways “constructivist,” that is to say, in some sense, they imply that what we ought to do is (at least in part) a function of what we (individually or collectively) say we ought to do (or would say we ought to do under specified circumstances).

IV.1 Moral Non-Cognitivism[15]

According to moral non-cognitivism, to say that A is obligatory is to express some sort of pro attitude towards A. If you have this pro attitude, you are inclined to do A. So, barring weakness of will, if you do not do A (i.e., if you are not even trying to do A) then you don't genuinely hold that A is obligatory. So, on this view, hypocrisy boils down to simple insincerity.

If Harriet tells John he ought to do A, and she fails to do A herself (and she's not a victim of weakness of will), the non-cognitivist will claim that Harriet doesn't really hold that A is morally obligatory. She is saying something she doesn't believe.[16] If John realizes this about Harriet, he can refuse to let her tell him to do A, on the grounds that she doesn't really hold A to be obligatory—as witnessed by her own failure to do A.

IV.2 Individual Ethical Relativism

Obviously there are different views that go by the name of individual ethical relativism. Let's take a simple one that will illustrate the point I am trying to make:

> An act is wrong if and only if the agent says it is wrong.[17]

(p.103) If you say that X is wrong and do it anyway, your act is wrong. This is the only way that an act can be wrong. So, hypocrisy would be the only thing that is always and necessarily wrong.

IV.3 Kant

Consider the universal law formulation of the categorical imperative:

> Act only according to that maxim whereby you can at the same time will that it should become a universal law.[18]

Since I am not a Kant expert, I won't presume to say what Kant meant by this, but let's take it in a fairly natural sense that many contemporary people do. As such, it seems designed to condemn hypocrisy. Someone who permits herself to steal or speed or fail to recycle, but makes efforts to discourage others from doing the same, is someone we judge to be a hypocrite and precisely, someone who is acting on maxims that she is not willing to allow others to live by. What would this view say about Harriet, Lindsay, and Moira? Now, of course, Kant (and any reasonable Kantian) would judge that some sets of principles are impossible for a rational person to will—universal cannibalism, for example. But a natural interpretation of Kant's idea does seem to allow for some discretion here.[19] So, for example, suppose that Harriet wants others to keep their children quiet in church so that she can pray in peace. She tells them to keep their children quiet; she frowns at them when they fail to do so, she complains to the priest, etcetera. But then when her children make a noise, she expects others to “suffer the little children” and to enjoy their youthful rambunctiousness. On the one hand, she wills it to be a universal law that parents keep their children quiet. On the other hand, she does not act on this maxim.

Kant seems to suggest that at least in some cases (perhaps this is one) the agent has a genuine moral option. It is permissible to let your kids make a noise—if you are willing to put up with the noise of others. And it is permissible to insist on quiet—if you are willing to do your part to ensure it. If so, then Kant (unlike M above) can rank the hypocritical Harriet below either the happily child-loving Lindsay and the strict disciplinarian, Moira. **(p.104)** This seems to accord the proper ranking to hypocrisy. In fact, it may seem that Kant's universal law makes this sort of hypocrisy or inconsistency the ultimate sin.

IV.4 Contractarianism

There are lots of different versions of this view, but the basic idea is that you ought to conform to rules that you have (or would have) agreed to in some contract.[20] While there are obviously limits to contracts reasonable people could agree to, it seems again that there are different workable contracts. We can see some of these played out in real contracts people make when forced to cooperate. Here are some options:

1. Members of a housing development agree to let each other paint their houses whatever color they want, have whatever ornaments, garbage, possessions in their yards, etcetera; or they all agree to make their homes aesthetically coordinated.
2. At the gym, each member agrees to wipe down the exercise machine after she has used it; or each member agrees to wipe down the exercise machine before she uses it, if she wants to.
3. Citizens of a town agree to share their books, toys, and tools—roughly in the manner of a public library—so, if someone needs a lawn mower, she can have the library recall it from whoever has it. While she uses it, she has to take care of it or pay a fine; or they agree to the sort of system of private property whereby they only get to use their own stuff and can do with it what they please—as long as it doesn't harm others.
4. Fellow citizens can agree to a system of democracy that chooses its leader by simple popular vote, or by the electoral-college system that we have in the United States.
5. They can all agree to help each other out in cases of need—taking on some burdens in exchange for some security against disaster; or they all agree to accept responsibility and risk for ourselves.

Given these options, it seems that Contractarianism will have resources to condemn as particularly objectionable someone who, on the one hand, expects her neighbors to live by one code while she herself insists on living by another. She is not playing the game, she is a free-rider.

(p.105) Contractarianism also has an elegant explanation for the retaliatory inconsistency that looked like a counter-example. It's permissible (and not hypocritical) to retaliate against the hypocrite, because she has broken her side of the contract. The contract between you and her is now null and void.

IV.5 The Golden Rule: Do As You Would Be Done By[21]

Although there are many difficulties with the interpretation of this rule, it seems again to be designed to rule out hypocrisy. For example, suppose that you want your house to be perfectly tidy, and expect your spouse to keep it that way. You should do your (equal) share to keep it perfectly tidy. The person who complains because her husband leaves a mess around and then does exactly the same thing is a hypocrite. The person who doesn't clean up the house but who doesn't complain is not violating the Golden Rule and neither is the person who does complain and does his share.

Things might get more complicated. What if he cares about a tidy house and she doesn't? Isn't it unfair to him? Well, perhaps this shows the flaws in the Golden Rule. But perhaps the Golden Rule can be raised to a sort of meta-level here. Suppose she cares about punctuality and he does not: while he cares about neatness and she does not. In that case, perhaps the Golden Rule would say to both of them: “Here's how you would be done by: you want your spouse to comply on the issue that you are obsessive about. You ought to do the same for her.” For example, it would say to the husband, “Either relax about keeping the house neat, or, if you're going to expect your wife to conform to your demands, you'd better start being punctual.” And it would say to the wife, “Either relax about punctuality, or if you're going to insist on punctuality, stop being such a slob.”

In other words, once again, we have a moral theory that ranks Harriet (the hypocrite) below either Lindsay (the slacker) or Moira (the consistent moralist). M-theorists tend to rank Harriet as better than either Lindsay or Moira. This can be treated as an argument for preferring “constructivist” theories, such as Kant (or the Kantian discussed here), Contractarianism, the Golden Rule, or Individual Ethical Relativism, as preferable to M-theories: **(p.106)**

1. Hypocrisy is a serious moral failing.
2. M-theories (moral realist where the facts about right and wrong don't depend on anyone's choices) don't imply that hypocrisy is a serious moral failing.
3. Constructivist theories imply that hypocrisy is a serious moral failing.

So, constructivist theories are prima facie better than M-theories.

V. Responses and an Account of What's Wrong with Hypocrisy

I earlier imagined four different possibilities regarding the action that the hypocrite was hypocritical about if M was true: A is morally obligatory, A is wrong, A is supererogatory, A is morally indifferent. On each, it seemed that M faced a difficulty that it would condemn a non-hypocrite more severely than the hypocrite for her treatment of A. Let's now consider a fifth possibility (that is perhaps consistent with any of the last four): suppose that M isn't clear about A; or we who are relying on M do not know what M says about A, or we and the hypocrite disagree about what the true morality (which we may or may not agree to be M) implies about A.

In that case, we (or John) can challenge Harriet like this: “Look, either A is, as you say, morally required or it is not. If it is, you are acting wrongly in not doing A. If it is not morally required, you are acting wrongly in pushing people to do P or in making them uncomfortable for not doing A. Either way, you are doing something wrong. You must see this. You are then knowingly acting wrongly.”[22]

(As I suggested above, John may not know that Harriet is a hypocrite. She may hide her inconsistency from him and others. In that case, imagine this conversation occurring in Harriet's own mind.) Such a speech could not be made to Moira or Lindsay. They can embrace the first and the second horn, respectively, and argue that they are not doing wrong.

Note that a significant part of the special badness here is the badness of knowingly doing wrong. Even if Harriet doesn't know exactly what it is that she is doing wrong, she knows (on this account) that she is doing something wrong. Moira, on the other hand, when she tells John to A and does so herself, is simply mistaken about what is morally required. Lindsay, similarly, is **(p.107)** mistaken about what is morally required. If we add, as I did in section III, that hypocrisy involves an inconsistency designed to place moral burdens on others while at the same time avoiding such burdens for oneself, then the objection becomes even more serious. Harriet is doing wrong, either in imposing burdens on others that they shouldn't have to endure, or she is avoiding taking a share of the burdens of morality that is legitimately hers. In this last respect, she is worse than an outright immoralist who sneers at morality.

It is worth noting that hypocrisy is primarily regarded as a vice, rather than a wrong action. If it is not axiomatic that a disposition knowingly to do wrong is a vice, then perhaps we need to consider different accounts of vice. Since this is not the place for more than a cursory survey of the matter, I shall rely on a very helpful review of recent work in virtue ethics, in which David Copp and David Sobel lay out the most prominent different accounts of virtue (and hence of vice) in the recent literature.[23]

Three consequentialist positions regarding virtue are as follows:

1. Virtues are states that tend to promote the good. On this view, vices would be states that tend to promote the bad, or tend to prevent the good.[24]
2. Virtues are intrinsically valuable. Vices are intrinsically bad.[25]
3. Virtues are states of loving and attempting to promote the good. Vices would be states of hating the good, or loving the bad, and attempting to frustrate the good, or promote the bad, or tending to neglect the good.

Michael Slote defines virtues simply as admirable states of character.[26] Rosalind Hursthouse considers virtues to be qualities that make one more likely to flourish or to be a good human being. Here she is echoing Aristotle and more recently Philippa Foot, who think of a good human being somewhat by analogy with a good horse or a good hammer. A good human being is good at doing what humans are in some sense designed to do. If we are talking about moral virtue, virtue cannot simply be the internal capacity to make oneself rich or powerful or healthy. Hursthouse must have a moralized notion of flourishing in mind.[27]

It seems clear that a tendency knowingly to do wrong would count as a vice by practically all of these different theories of virtue. Notice that the theories sketched here are all rather schematic. They don't commit themselves to any particular account of value or right action. Driver's account of **(p.108)** virtue as an internal characteristic that tends to promote value is consistent with hedonistic, preferentist, and perfectionist versions of consequentialism, for example. It seems clear, though, that an agent who had a tendency knowingly to do wrong would have a tendency to promote disvalue. It is possible that she would occasionally produce more value than disvalue, but the expectation is that this would probably not be the case in most circumstances.

Similarly, a view according to which virtue involves a love of the good and a desire to act to promote it, and vice is the opposite, would hold that the hypocrite has the wrong conative attitude towards the good, and hence, is guilty of a vice.

Views like Hurka's and Slote's that treat vice as intrinsically bad or contemptible (if that's the opposite of admirable) would probably concur that hypocrisy is vice, since it is almost universally regarded with revulsion. Theories like Hursthouse's and Foot's that treat virtue as a matter of flourishing or functioning well qua human being would probably regard hypocrisy as representing the sort of inner conflict that Plato regarded as a lack of flourishing, if not the ultimate form of self-destruction.

If this is correct, then hypocrisy is almost certainly a vice—although we have different accounts of why it is objectionable.

VI. Conclusion

I have argued that hypocrisy is an inconsistency between one's avowed or apparent values and one's behavior, where this inconsistency has the effect or is intended to have the effect of imposing morality's burdens on others while avoiding them for oneself.[28] Like many such terms, hypocrisy is undoubtedly rather vague so there might be some borderline cases. I have argued that this is a vice because it involves an agent knowingly doing wrong (in at least part of her mind) and involves unfairness with respect to others. Note that this view is similar to a view expressed by Roger Crisp and Christopher Cowton, that what is especially objectionable about the hypocrite is his lack of moral seriousness.[29]

(p.109) References

Bibliography references:

Boucher, D. and Kelly, P. (eds.) (1994), *The Social Contract from Hobbes to Rawls* (New York: Routledge).

Copp, D. and Sobel, D. (2004), “Morality and Virtue: An Assessment of Some Recent Work in Virtue Ethics,” *Ethics* 114 (April): 514–54.

Crisp, R. and Cowton, C. (1994), “Hypocrisy and Moral Seriousness,” *American Philosophical Quarterly* 31(4): 343–9.

Driver, J. (2001), *Uneasy Virtue* (Cambridge: Cambridge University Press).

Flanagan, C. (2007), “No Girlfriend of Mine,” *The Atlantic Monthly* November: 139–44.

Foot, P. (2001), *Natural Goodness* (Oxford: Oxford University Press).

Gauthier, D. (1986), *Morals by Agreement* (Oxford: Oxford University Press).

Hare, R. M. (1963), *Freedom and Reason* (Oxford: Oxford University Press), 67.

Hurka, T. (2000), *Virtue, Vice, and Value* (Oxford: Oxford University Press).

Hursthouse, R. (1999), *On Virtue Ethics* (Oxford: Oxford University Press).

Kant, I. (1887), “The Metaphysic of Morals,” in *The Cambridge Edition of the Works of Immanuel Kant* (1996 vol. on Practical Philosophy), ed. and trs. Mary J. Gregor (Cambridge: Cambridge University Press), 474.

—— [1785] (1981), *Grounding for the Metaphysics of Morals*, trs. James W. Ellington (Indianapolis, IN: Hackett), 421.

Parfit, D. (forthcoming), *On What Matters* (Oxford: Oxford University Press).

Slote, M. (2001), *Morals from Motives* (Oxford: Oxford University Press).

Szabados, B. and Soifer, E. (1999), “Hypocrisy, Change of Mind, and Weakness of Will: How to do Moral Philosophy with Examples,” *Metaphilosophy* 30(1/2, January/April): 71.

Turner, D. (1990), “Hypocrisy,” *Metaphilosophy* 21(3, July): 262–9.

Van Roojen, M. (2009), “Moral Cognitivism vs. Non-Cognitivism,” *The Stanford Encyclopedia of Philosophy* (Fall), Edward N. Zalta(ed.), forthcoming URL ⟨http://plato.stanford.edu/archives/fall2009/entries/moral-cognitivism/⟩.

Notes:

(1) *The Holy Bible*, New International Version. (All Biblical references below are to this version.)

(2) *The New York Times*, Opinion Page, 20 October 2007.

(3) Flanagan (November 2007).

(4) Crisp and Cowton (1994: 31.4: 343–9).

(5) Szabados and Soifer (January/April 1999) make this point.

(6) Turner (July 1990). I wouldn't necessarily agree with all the examples Turner cites—e.g., beliefs vs desires—but the idea of an inconsistency between apparent moral attitudes and behavior seems right.

(7) Szabados and Soifer (January/April 1999).

(8) R. M. Hare makes this point in *Freedom and Reason* (1963: 67).

(9) Weakness of will is a terribly complex phenomenon. I can't hope to do justice to the psychological mechanisms involved. Hopefully, however, this use of excluded middle to cover all cases of weakness of will won't be objectionable. In case it is, I'll discuss borderline cases below.

(10) Hypocrisy seems more problematic when the criticism, blame, condemnation of others is more harsh. A parent encouraging her children to avoid drugs is expressing concern and compassion, not revulsion and indignation, and hence, this doesn't seem to count as hypocritical—perhaps there's a hint there about what's problematic about hypocrisy. It's not simply inconsistency, it's unfairness—placing burdens on someone else that one is not willing to bear oneself. The burdens are for some that burden of living up to the moral law—which may be onerous, and for others, the burden of suffering punishment or moral opprobrium.

(11) Bela Szabados and Eldon Soifer (January/April 1999: 63–6) discuss a case of someone whose aesthetic judgment changes over time and who seems to exemplify hypocrisy.

---

(12) Kant (1887).

(13) Worry: Maybe it isn't Harriet's business to see to it that John be doing what he ought. Maybe it's even wrong for Harriet to do so. But if M says this, then presumably, M should say the same to Moira.

(14) Some M-theorists may regard this as unproblematic. I've often seen or heard utilitarians argue that members of their audience ought to avoid meat, or give 50 per cent of their incomes to charity or give a kidney to a stranger. When challenged with the objection that they themselves do not do these things, they reply that this objection is an irrelevant ad hominem. If their argument to the conclusion that one ought to give a kidney to a stranger is sound, its conclusion is true and would be true whether it were offered by a saint or a scoundrel. While this is true, it skirts dangerously close to treating moral issues as mere intellectual issues that do not have any impact on actual behavior. It is worth noting that such utilitarians typically do not believe that those who fail to live up to their very high standards deserve punishment, blame, or guilt for their failures. (See my “Elbow Room for Consequentialists,” *Analysis* October 1992). It is also worth noting that such utilitarians typically reject moral internalism. A different sort of response an audience member could make would be to say, “Yes, you're right. I ought to give up meat, give away most of my money, and donate a kidney to a stranger. But, like you, I'm not going to do it, and I'm not going to feel guilty about it.”

(15) For a helpful outline of non-cognitivism, see van Roojen (2009).

(16) I use the term “believe” here somewhat loosely. Moral non-cognitivists hold that technically moral attitudes such as “Stealing is wrong” are not beliefs, but conative attitudes such as wishes.

(17) Don't worry too much about the circularity in this definition.

(18) Kant ([1785] 1981: 421).

(19) My colleague Hud Hudson, who knows Kant's work far better than I do, claims that Kant himself did not believe in this sort of discretion. Any two perfectly rational wills would converge on the same set of moral rules. If so, true Kantianism is more like one of the M-theories I describe above.

(20) See Boucher and Kelly (1994) and Gauthier (1986).

(21) For an interesting discussion of the Golden Rule and its relationship to Kant's Categorical Imperative, see Parfit, *On What Matters* (forthcoming).

(22) Notice that this speech works just as well in cases where A is straightforwardly wrong or straightforwardly morally obligatory, etcetera.

(23) Copp and Sobel (April 2004).

(24) Driver (2001).

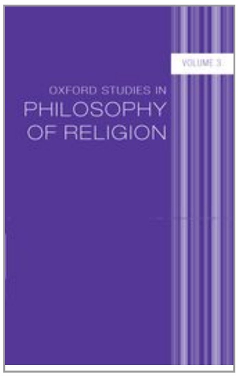
(25) Hurka (2000).

(26) Slote (2001).

(27) Hursthouse (1999) and Foot (2001).

(28) The phrase “morality's burdens” here is slightly loose, meaning roughly the burden of doing what one believes to be right or of paying the consequences for not doing what one believes to be right.

(29) See Crisp and Cowton (1994: 347): “All of these failures [different kinds of hypocrisy] would be avoided by a person who took (morally) seriously the relationship between her self and morality, someone who demonstrated an active and genuine concern to be moral. This concern is in a sense a metavirtue, since it is a concern to be virtuous. But it is nevertheless an excellence, the lack of which in various ways is likely to issue in hypocrisy when its possessor is engaging in moral practice or the moral language game. This also explains much of what is bad about hypocrisy. If anything is morally blameworthy, then lack of concern for morality itself surely is.” I am grateful to an audience at Calvin College, a class at Western Washington University, Hud Hudson, Daniel Howard-Snyder, Mary Howard, and an anonymous referee for this volume, for helpful discussion of the issues in this paper. I am also grateful to Western Washington University for professional leave that enabled me to work on this project.



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

The Argument from Consciousness Revisited

Kevin Kimble
Timothy O'Connor (Contributor Webpage)

DOI:10.1093/acprof:oso/9780199603213.003.0007

Abstract and Keywords

The argument from consciousness maintains that correlations between brain states and conscious states of persons require explanation but cannot be given an adequate scientific explanation. The chapter then argues that the best explanation of these correlations is that they are the result of the work of a purposeful supernatural agent. The aim is two-fold. First, the chapter considers and rebuts recent attempts in the philosophy of mind to defend a physicalist account of the phenomenal character of experience (accounts which, if successful, would undercut the core premise of the argument from consciousness). It then considers two versions of the argument from consciousness and contend that they are defective, since they overlook a naturalistic form of explanation that is available even on a robustly dualistic account of conscious states. However, it goes on to show that the argument may more plausibly be recast by treating the very form of explanation of conscious states it outlines as a further datum in the fine-tuning version of the design argument.

Keywords: consciousness, dualism, emergence, fine-tuning, naturalism, phenomenal concept, physicalism, qualia, theism

More than two decades ago, Richard Swinburne and Robert Adams put forth an argument for theism that they aptly labeled the *argument from consciousness*.[1] A thumbnail sketch of the argument goes like this: There are facts involving correlations between brain states and conscious states of persons for which rational inquiry demands a satisfying explanation. There are but two broad forms such a possible explanation may take: the correlations can be explained either through more basic scientific laws or by the intentions and actions of a powerful personal agent. Since the correlations' facts cannot be given an adequate scientific explanation, the best explanation is that they are the result of the work of a purposeful agent.

Our aim in what follows is twofold. First, we consider sophisticated recent attempts in the philosophy of mind to defend a robustly physicalist account of the phenomenal character of experience, accounts that, if successful, would undercut the core premise of the argument from consciousness (AC). We will try to show, however, that these accounts fail. We then consider the version of AC advanced by Swinburne and Adams. We contend that their versions are defective, since they overlook a naturalistic form of explanation that is available even on a robustly dualistic view of conscious states. However, we go on to show that the argument may more plausibly be recast by treating the very form of explanation of conscious states we outline as a further datum in the currently popular fine-tuning version of the design argument. We do not attempt to determine whether the fine-tuning argument is ultimately successful.

(p.111) I. Direct Phenomenal Awareness

The fact that there are strong correlations between underlying neuro-physical states of the brain and conscious mental states is not in serious doubt. Philosophers of mind commonly divide the latter into two broad classes, intentional states and phenomenal states. The former has commonly been equated with propositional attitude states (beliefs, desires, and so forth), while the latter has been thought to consist of the “raw feels” of experience—the qualitative way things look, sound, taste, and so on. Although we seriously question this bifurcation of mental states into these two distinct kinds, for our present purposes we may safely focus solely on the feature of phenomenal character.[2] Although the present state of neuroscience does not yield precise hypotheses, it is highly plausible that certain neuro-physiological types directly correlate with certain phenomenal types. So we may suppose that brain state R correlates with a phenomenally reddish look, brain state B correlates with awareness of phenomenal blue; T correlates with tangy orange sensations, P with a certain piercing auditory quality, and so on. While these cartoonish schema certainly oversimplify how mind-brain correlations actually work, and the way in which brain states and phenomenal states map onto each other, we can assume something like this for the sake of simplicity and concreteness.

Physical-phenomenal correlations of this kind can count as evidence for design only given an assumption that phenomenal properties are *ontologically distinct* from their physical-functional correlates. Unless the most basic phenomenal properties are ontologically novel features of the world, fundamental features over and above the features characterized by physical theory, the argument that the correlations constitute a credible design-type fact to be included among other design evidence will be open to a straightforward refutation. Identity is the most satisfying explanation of the correlations that is imaginable, and hence one that cannot be improved in theological terms.

Thus the proponent of the argument from consciousness needs to make plausible that phenomenal qualities themselves are ontologically primitive. Swinburne and Adams devote little space to this question, and the discussion they do carry on is dated. In particular, there are two prominent materialist accounts of phenomenal qualities in recent philosophy of mind, unanticipated by Swinburne and Adams, that we will consider here.

First, some contemporary thinkers insist that the connection between certain physical and phenomenal properties is one of identity, despite the **(p.112)** fact that we cannot grasp the connection a priori. They point out that we may conceive of a single property in more than one way, corresponding to different modes of epistemic access to it. In the case of physical/phenomenal features, we have two wholly distinct sets of concepts corresponding to the modes of scientific theoretical identification and introspective awareness. These thinkers then argue that, due to the peculiar nature of our *phenomenal concepts*, the fact that we cannot discern the physical-phenomenal identity via a priori reflection is unsurprising, and it does not follow that the properties we variously conceive of are ontologically distinct.[3] We shall call this strategy *type physicalism*.

Second, a very different strategy for identifying phenomenal qualities with physical properties has developed in recent years, motivated by a view of phenomenal experience called *representationalism*. Representationalism is the general term for a family of related positions, but the core idea among physicalist versions of the theory is that phenomenal qualities are to be identified with certain externally constituted representational properties of experience, rather than with internal states (physical-functional or primitively mental) of the experiencing subject.[4] Phenomenal character is identical to an experiential state's intentional content which, depending on the particular view, consists of entities such as external properties, propositions, or states of affairs. On plausible physicalist assumptions, such content involves structured physical properties or dispositions, such as (for colors) the disposition of a surface to reflect a certain percentage of light from the visible spectrum. Thus, on representationalism, the look of the redness of the rose to a subject is none other than the external physical property of redness. We will call physicalistically acceptable representational properties *p-representational* properties. Since phenomenal qualities are extra-mental features, there is no primitive mind-body correlation to be explained.[5]

Given the growing popularity of these a posteriori physicalist accounts of the phenomenal character of experience, a defender of the argument from **(p.113)** consciousness must make the case that all such strategies are untenable. We will now undertake that task.

The case for the ontological distinctness of phenomenal qualities from any physical-functional properties is broadly introspective. When I look at a luscious mandarin orange in bright sunlight and bite into it, at the same time focusing my attention on the qualitative way I experience it, I become *phenomenally aware* of an orange-ish quality presented to me in a sensate visual way—a smooth and uniform, simple feature impressing itself upon and occupying a region of my visual field. In addition, I am aware of a sweet tangy sensation on my tongue, which is, once again, a fairly simple and uniform taste quality. These contents of my awareness are “transparent” to me in the sense that I have unmediated access to what it is like to experience those qualities. And I recognize certain facts about those qualities that bear on the a posteriori identity thesis. To an approximation, let us say that a property is non-structural if and only if its instantiation does not even partly consist in the instantiation of a plurality of more basic properties by either the entity itself or its parts.[6] Now, my phenomenal visual field may be highly complicated, consisting of differently sized, shaped, and colored objects arrayed in a certain manner. But I can focus on sufficiently small subregions of the field that feature nonstructural or simple properties, such as a determinate (phenomenal) shade of blue. Quite generally, it seems that the contents of phenomenal awareness ultimately decompose into such simple, non-structural features, all of which are themselves objects of awareness.

Our simple argument from direct awareness goes like this: phenomenal awareness reveals an intrinsic simplicity in the elements comprising the qualitative character of my experience. However, the physical-functional (or p-representational) properties with which that character is supposed to be identified do not possess this intrinsic simplicity. Type-physicalists identify phenomenal character with complex neuro-scientific properties, while representationalists claim that phenomenal qualities are externally constituted low-level physical properties. Both candidates are heterogeneous structural features. Hence, the immediate contents of my phenomenal awareness, the phenomenal character and qualities of which I am directly aware, are not identical to underlying physical-functional (or p-representational) properties.

Moreover, in my phenomenal awareness of the orange-ish look of the mandarin orange, there must be some explanation for the fact that the phenomenal property I directly grasp is not presented to me either as a neuro-physical state of some sort, or as a surface reflectance or some other underlying physical **(p. 114)** property, as a physical color would be in its own nature. The advocate of ontologically primitive qualia has at his disposal a ready explanation for this. Phenomenal properties are to be identified with neither internal physical-functional states of the brain nor extra-mental p-represenitional properties. By appealing to the notion of immediate awareness or *acquaintance*, the qualia theorist can make the case that the qualities one has direct access to are intrinsic, non-structural features of experience. This explanation provides adequate grounds for our immediate awareness of phenomenal character and best accounts for its intrinsic simplicity vis-à-vis the objects and content of that awareness.

Now we can deploy this argument from direct access in the task of effectively debunking recent attempts to defend the a posteriori identity thesis. It is important to notice that the argument we advance here does not make the assumption that every distinct concept expresses a distinct and genuine property, nor does it make use of dubious premises concerning the relationship between conceivability and possibility. Hence it is to be sharply distinguished from other widely known and discussed arguments for property dualism.[7]

One physicalist response to our argument contends that what appears to be a phenomenal/physical property distinction is in reality merely a distinctness between modes of presentation of the selfsame property. However, the problem just resurfaces as a question about the nature of the relevant modes of presentation. Suppose that physical blueness is a certain surface reflectance property B, and that what I am phenomenally aware of—the particular visual qualitative character that is presented to my awareness—is a mode of presentation of B.[8] As the direct object of my awareness, this mode of presentation M is a feature or property of either B itself or one of my representations of B, and thus M stands in need of some acceptable physicalistic explanation or reduction so as to avoid an undesirable “dualism” of properties.

Materialists must deny that awareness of phenomenal properties and content involves any distinct mode of presentation of the referent property or content. But how can there be a *phenomenal presenting* of a property or content to the subject's awareness unless there is also some mode of presentation **(p.115)** involved? The answer, for the a posteriori physicalist, is that the referent property or property-involving content presents itself directly to awareness; the phenomenal property is revealed directly as it is in itself, and the subject grasps this property via the exercise of an appropriate *phenomenal concept*. It is the peculiar nature of these phenomenal concepts that accounts for why we don't grasp phenomenal qualities as physical-functional properties. There are various ways this idea has been elaborated.

Theories of phenomenal concepts fall neatly into two broad classes: causal-recognitional accounts and self-referential or quotational accounts. On the first type of account, phenomenal concepts are a special kind of recognitional concept that pick out phenomenal qualities directly by demonstrating or naming them, without relying on theoretical or background information. Loar, McLaughlin, and Tye each hold this sort of view.[9] Loar's and McLaughlin's proposals are virtually identical, while Tye's view differs from theirs crucially on the issue of the nature of the content of phenomenal concepts. Loar and McLaughlin hold that the referents of phenomenal concepts are internal neuro-biological properties, whereas Tye holds that their referents are externally individuated representational contents.

The self-referential or quotational account is similar to the first type of view except that the phenomenal qualities attended to in experience serve as constituents of the very phenomenal concepts that pick them out, thus conferring upon them a self-referring function. The basic idea is that there is a subset of neural states that are constitutive of phenomenal states (or serve as the core realizers of such states), and a given phenomenal state may get embedded in a more complex representation by a sort of mental "quotation" process which incorporates the original state as a constituent. The representation in turn takes on a sort of reflexive indexical function of referring to the very neural (phenomenal) state which in large measure constitutes the representation itself. The representation is a phenomenal concept. Block, Balog, and Papineau all propose versions of this type of account. While the details of these individual accounts differ in various places, they are all of one accord in their implicit answer to the problem of intrinsic simplicity.[10]

In rough outline, here is how the story goes as concerns our access to and awareness of phenomenal properties. The subject of an experience, in **(p.116)** exercising an appropriate phenomenal concept during an act of introspective awareness, directly grasps the phenomenal property *as it is in itself*. Type-physicalists such as McLaughlin speak of phenomenal qualities as constituting their own modes of presentation for the subject, thus revealing themselves to the subject's awareness as they are in and of themselves—yet not by grasping those properties via any essential or non-essential aspects of them. Phenomenal concepts lack any descriptive content whatsoever, functioning much like non-descriptive name concepts or type-demonstrative concepts, and hence do not reveal anything about the essential nature of phenomenal properties.[11]

Other theorists write as if they deny any mode of presentation at all. Tye, for example, in discussing his view on color, says “Indeed, on my view, colors are not presented to us in sensory experience under any mode of presentation at all. Our awareness is direct.”[12] This is true of first-order phenomenal awareness as well as introspective awareness of phenomenal content. With respect to the latter, Tye claims that introspecting a perceptual state involves exercising one or more phenomenal concepts that pick out the represented content directly and refer via the direct causal connection they have with their referents.[13] In agreement with McLaughlin, he holds that these concepts have no associated reference-fixers and contain no descriptive content at all. Rather, “the referent is presented without the assistance of associated features distinct from the referent which the thinker a priori associates with it. There is no separate guise that the referent takes in the thinker's thoughts.”[14] A subject knows that she is in a visual state “as of” seeing blue just by attending to how things look to her, not by attending to something else connected with it. What the phenomenal concept directly picks out is the complex representational content, the key constituent of which is a surface reflectance-involving state of affairs. The subject is phenomenally aware of the state of affairs that has a surface reflectance *b* as a component.

Despite some differences over how they use terminology, Tye and McLaughlin are in basic agreement on the issue of modes of presentation of **(p.117)** phenomenal properties. When Tye says that a phenomenal property is not presented under any mode of presentation, he means that there is no aspect or property separate from the referent itself that serves to anchor or ground application of the relevant phenomenal concept. Since he apparently accepts the fact that the referent property is “phenomenally presented” to the subject (e.g., there is a bluish look or feel in my visual field), the phenomenal concept picks out the property directly and as it really is, under no other guise. This seems to be equivalent to McLaughlin's claim that phenomenal qualities serve as their own modes of presentation.[15] Of course, as noted previously, the *objects* of phenomenal awareness are radically different on the two views. McLaughlin holds that phenomenal character is constituted by certain neuro-scientific properties of the brain, whereas Tye holds that qualitative character consists of a complex representational state, a crucial component of which are certain low-level physical (surface) properties in the extra-mental environment.

II. An Inconclusive Argument Against a Posteriori Physicalist Accounts of Direct Phenomenal Awareness

The account of direct access to phenomenal properties that we have offered encompasses claims that stand in tension with the key identity claim made by proponents of a posteriori physicalism. Although we might try to make explicit this tension by formulating these claims in terms of phenomenal concepts, the problem is independent of theories of phenomenal concepts, in that it arises for anyone who accepts a form of “direct access” akin to what we have been calling phenomenal awareness—regardless of whether or not such awareness is mediated by the exercise of phenomenal concepts. The following locutions, which we will use somewhat interchangeably, are intended to capture the sense of “direct access” to phenomenal qualities that physicalists such as McLaughlin and Tye allow that we have. They are also intended to be broad enough to permit a form of direct phenomenal access that does not require concept deployment. (If non-conceptual awareness of **(p.118)** phenomenal qualities seems inconceivable to you, then no harm will be done by reading the following locutions as claims about our grasp of phenomenal properties via the exercise of phenomenal concepts.) When we say that a subject S *directly accesses* a phenomenal property Q in undergoing a perceptual experience, we mean roughly the following: S is phenomenally aware (“*q-aware*,” as distinct from “*p[perceptually]-aware*”) of Q; or, S phenomenally grasps Q in an unmediated way; or, Q is directly revealed or presented to S in an act of phenomenal or introspective awareness. The form of awareness in question might be first-order or higher-order, whereby S directly grasps phenomenal quality Q and conceives it by deploying a phenomenal concept C in an act of introspection.

Now, the following claims made by a posteriori physicalist theories of phenomenal access appear to be in tension:

i. *S directly accesses Q as it truly is in itself*. Q is revealed to us as it really is, leaving no room for illusion. For example, we are not *q*-aware of phenomenal blue under any distinct mode of presentation or manifestation of it; rather, phenomenal blue is presented to our awareness as it is in itself—meaning that the property (or referent of the appropriate phenomenal concept) itself serves as its own mode of presentation.
ii. *Q is a physical-functional property*. The object of *q*-awareness (Q) is an underlying physical property—either a neuro-scientific property within the agent or a micro-physical surface property in the external world. The nature or essence of phenomenal blue consists of either a complex neural state of the visual cortex (type-physicalism) or a particular surface-reflectance disposition of an external-facing surface (representationalism).
iii. *S does not directly access Q* *as* *a physical-functional property*. We are not *q*-aware of phenomenal qualities as physical-functional properties of any sort; they are not presented to awareness as physical-functional properties.

Claims (i) and (iii) seem jointly to preclude claim (ii): if phenomenal qualities are revealed to us in experience *as they really are*, and if we are not aware of them *as* underlying physical-functional properties, then how can they be identical with such properties? Alternately, if we directly access phenomenal properties as they truly are, and if they are just identical to certain low-level physical-functional properties, then it seems that we should be phenomenally aware of their underlying physical-functional nature.

Thus, there appears to be a straightforward argument against physicalism about phenomenal character. Indeed, this reasoning parallels a similar argument against a posteriori type-physicalism offered by Horgan and Tienson **(p.119)** (2001). In circumstances in which a subject S is occurrently aware of a phenomenal quality Q while undergoing a perceptual experience, the argument claims the following:

1. When S is phenomenally aware (Q-aware) of phenomenal quality Q, S is aware of Q as it truly is in itself.
2. When S is q-aware of Q, S is not aware of Q as a physical-functional property.
3. If S is aware of Q as it is in itself and S is not aware of Q as a physical-functional property, then Q is not a physical-functional property.
4. Conclusion: Q is not a physical-functional property.[16]

Is this apparently straightforward argument against physicalism in fact sound? No; there is a flaw in the argument that occurs at step (3). The conditional expressed by (3) is false. It simply does not follow that if S is aware of Q as it is in itself and S is not aware of Q as a physical-functional property, then S is aware of Q as a *non*-physical-functional property. And this intermediate step is required if the conditional expressed by (3) is to have any plausibility. From the fact that S does not directly access Q as an F, it does not follow that S directly accesses Q as a non-F. In general there is a distinction to be drawn between "*not* being aware of *a* as a B" on the one hand and "being aware of *a* as a *non*-B" on the other. A child's not being aware that the figure she is looking at is an octagon does not entail that she is aware of the figure as a non-octagon. And her awareness may nevertheless be considered a kind of direct awareness of the figure as it is in itself. (A similar point can be made about awareness of certain phenomenal qualities. I may not be aware of a bluish look in my visual field as a property *of my experience*, but that does not mean that I am aware of the bluish look as being a property *of something other than my experience*.) In his response to Horgan and Tienson, McLaughlin appeals to a similar distinction, claiming that in order for E to be revealed as an F, the subject must bring E under the concept of F.[17] He points out that we do not grasp phenomenal qualities *as* physical-functional properties for the simple reason that in grasping the relevant quality via some phenomenal concept C, we do not bring that quality under any physical functional concept. But it does not thereby follow that the phenomenal quality is *otherwise than* a physical-functional property.

(p.120) In accordance with (ii) above, physicalists maintain that Q is a physical-functional property. Moreover, implicit in (ii) is the idea that the nature or essence of Q is *exclusively* physical-functional. Denying this would be tantamount to conceding that phenomenal properties have dual (or multiple) essences, one of which is non-physical-functional, and this in turn would undercut physicalism in favor of a kind of dual-aspect theory. Thus, physicalists such as Tye and McLaughlin maintain that the essence of phenomenal properties is wholly physical-functional.

Tye and McLaughlin have successfully dodged Horgan and Tienson's argument. Whether their response to it undermines our original argument from direct awareness is something we consider in due course. Let us first observe that we still seek a constructive answer to the following question: given that we directly access a phenomenal quality Q as it really is in itself, and given that the nature of Q is exclusively physical-functional, then why *don't* we directly access Q as a physical-functional property? If the direct access we are talking about involves conceptualization of Q, then on one level the answer is that we simply don't access Q via the exercise of a physical-functional concept. But if Q's very physical-functional nature is directly revealed to our awareness, then why doesn't our immediate grasp of Q involve, at least in part, the exercise of an appropriate kind of physical-functional concept to pick it out?

Tye and McLaughlin's answer is to deny the antecedent in the foregoing question: the *essence* of a phenomenal property is *not* revealed to our awareness, and hence we do not directly access that essence.[18] We directly grasp the phenomenal property as it is in itself—the property presents itself as it really is—but this does not mean that we directly access its nature or essence (even in part). And the fact that we don't directly access phenomenal properties *as* physical-functional properties implies that we don't access their nature(s), since they are exclusively physical-functional. Thus, in addition to (i)–(iii), physicalists of the Tye-McLaughlin stripe also hold:

iv. S does not directly access Q's nature or essence.

The way McLaughlin and Tye resolve the apparent tension in our initial claims (i)–(iii) listed above is by offering a further elaboration on the relationship between (i) and (iii), which involves the additional claim (iv). (iv) qualifies (i) and helps explain why (iii) is true. The fact that we don't directly **(p.121)** grasp the essential nature of phenomenal properties explains why we don't conceive of them physical-functionally. We may think of (iv) as a denial of the *Doctrine of Revelation*:[19]

> *Revelation (R)*: The intrinsic nature of a phenomenal property Q is at least partly revealed to us in our *q*-awareness of the property.

The remaining task for the a posteriori physicalist is to offer a plausible story that explains or clarifies the consistency between (i) and (iv):

i. S directly accesses Q as it truly is in itself.
iv. S does not directly access Q's nature or essence.

In summary, here is what Tye and McLaughlin say on the matter. Our unmediated grasp of phenomenal properties is attained by exercising non-physical-functional concepts that refer directly to those properties. We are directly aware of Q as it is in itself *because* our phenomenal grasp of Q involves no distinct mode of presentation of Q, but a direct demonstration of Q itself. That is, the phenomenal concept by means of which we access Q demonstrates and presents Q directly. So in presenting itself directly to our awareness, Q serves as its own mode of presentation. This is the needed clarification of (i). At the same time, our direct grasp of Q is not by means of some description or mode of presentation of Q that captures or is an aspect of it, or part of its essence, a description or mode that would allow a conceptualization of it as such and such. Our access to Q does not involve fixing reference via some associated description of Q, since the phenomenal concept that affords a direct grasp or access of Q contains no descriptive content. The phenomenal concept functions as a naming device or type-demonstrative, and hence does not conceptually reveal anything about the essential nature of Q. So we are not phenomenally aware of Q as a physical-functional property, since we do not access Q via a physical-functional concept containing the appropriate description. This provides the required explanation for (iv). Thus according to Tye and McLaughlin, we can be directly aware of phenomenal properties as they truly are in themselves, without illusion, and yet without grasping their natures. This explanation renders mutually consistent the central claims (i)–(iv) put forth in their account of direct access to phenomenal character and content.

(p.122) III. The Content of Phenomenal Concepts

We contend that the Tye-McLaughlin position on phenomenal access and awareness is unstable, in that their position when pressed will either give way to primitivism concerning phenomenal properties, in accordance with the thesis of intrinsic simplicity, or lead to a denial of the phenomenological data as we encounter it in experience. In the end, respecting the phenomenological evidence requires embracing a limited-revelation thesis, and this picture fits best with a primitivist view of the nature of qualitative character.

Let's consider (i). If we press for greater clarity on issue of what it means for S to directly access Q “truly as it is in itself,” Tye and McLaughlin claim that Q presents itself as it is, as its own mode of presentation. On the face of it, this claim appears to be affirming a *genuine presentation* of Q to S's phenomenal awareness; and since Tye and McLaughlin are phenomenal realists, they take the phenomenology of experience rather seriously. If a phenomenal quality is visually presented to me (e.g., if I directly access it through the visual modality by application of a phenomenal concept appropriate to that modality), then the quality is presented to me in a certain way, or as something. That “certain way” or “as something” is the way it looks to me. True, the mode of presentation just is the referent property itself, but it is nevertheless a substantive presentation of that property. Consider, for example, the experience of staring at the sparkling bluish appearance of the Pacific Ocean on a bright afternoon. As one attends to the bluish character of one's visual field, one's phenomenal grasp of the bluish quality constitutes a particular way things seem to one, and that way *is* the direct presenting of the phenomenal property itself, i.e. its constituting its own mode of presentation. This much is clearly present to one in the experience, and once we have come this far, we can go a step further. The presentation of the phenomenal quality is something one has a substantive and determinate grasp of, gleaned from the visual phenomenology itself. This substantive grasp reveals primitive, non-structural features as the immediate contents of one's *q*-awareness. At least that is the way things clearly *seem* to be, phenomenally speaking. The bluish look of the ocean is intrinsically simple in the following sense: what one directly accesses (what presents itself to one's awareness) seems to be a smooth, uniform, homogeneous quality, one that is not further dissectible into parts or discernible patterns, as far as the actual phenomenology goes. We have a substantive and determinate conception or grasp of its simple, non-structural qualitative nature. We claim that this feature of *intrinsic simplicity* thoroughly pervades the phenomenology of **(p.123)** perceptual experience in all of the basic sense modalities. This important aspect of the phenomenal character of experience we can call the *Intrinsic Simplicity condition* (IS):

> *Intrinsic Simplicity (IS)*: Phenomenal character and content is constituted (at least in part) by intrinsic, non-structural features directly accessed in phenomenal awareness.

IS holds for phenomenal qualities grasped via any of the standard sense modalities. The sweet, tangy taste of a mandarin orange when you bite into it phenomenally presents itself as a simple, uniform gustatory sensation of a particular kind; a vocal note sung by a choir has a distinctive auditory quality that can be described as a smooth, homogeneous, intrinsic feel impinging on one's awareness. Examples of such intrinsic, qualitative sensations readily come to mind for the other sense modalities as well.

If what we have said so far is on the right track, then in phenomenal and introspective awareness it looks like we do have a positive, substantive grasp of the objects of that awareness, phenomenal properties. Perceptual phenomenology reveals an awareness of primitive, non-structural features, in accordance with *IS*. This in turn implies that we have a substantive grasp of *the way a phenomenal property really is*, which means a grasp of (at least some aspect of) its real nature or essence. Let's call this claim *minimal* or *partial descriptivism*. (We insert the “at least some aspect” qualifier in order not to preclude a metaphysical picture on which phenomenal states have dual aspects, primitively phenomenal and physical.)

In introspecting a phenomenal property, either we grasp the property by applying concepts that involve some descriptive content, or we grasp the property via a phenomenal concept, which involves no descriptive content whatsoever. Tye and McLaughlin deny that phenomenal concepts carry any descriptive content or that their application involves conceptualization of the relevant phenomenal quality “as such and such.” Phenomenal concepts in their view are purely recognitional, and so their exercise does not involve the application of other concepts that would allow the subject to conceive of the quality or fix reference to the quality via a description (including any physical-functional description). But Tye and McLaughlin further assume that any recognitional-phenomenal concept, since it is not descriptive *in that way*, thus contains no substantive content at all. Here it seems that Tye and McLaughlin present us with a forced choice between two extreme claims: a phenomenal concept either grasps the property it refers to via a *full* descriptive content that characterizes the property in question or else it grasps the phenomenal property in no substantive way at all. We claim that this amounts to a false dichotomy, because there is a third possibility—there **(p.124)** are recognitional-phenomenal concepts whose exercise does not involve the use of other concepts but which are nonetheless substantive in that they ground the application of further concepts which do partially and accurately characterize what the phenomenal property is like.

Thus we contend the phenomenological data suggest that the nature of the content of immediate phenomenal experience lies in between these two implausible extremes—that we possess a robust and determinate grasp of phenomenal qualities (which amounts to more than mere naming or demonstration) but at the same time one that falls short of a full description or conceptualization. We do not directly conceive of or access phenomenal properties as instances of full-fledged, determinate property types of one sort or another (micro-physical, neuro-scientific, or otherwise). Nevertheless, we directly grasp them in a substantive way that allows us to form rudimentary concepts which truly characterize those properties, and these concepts can be seen on reflection to be incompatible with their being identified with the sorts of highly structured physical properties or relational functional properties recognized by physics and neuroscience.

To elaborate: the phenomenology of perceptual experience is not completely neutral concerning the nature of the properties that we access in phenomenal awareness. The nature of phenomenal experience/the content of phenomenal concepts reveal both the presence of certain features and the absence of other features, which suggests that phenomenal character is not to be identified with underlying physical properties. In the examples considered above, we identified phenomenal color, taste, and sound qualities, all of which exhibit the positive feature of intrinsic simplicity as described in *IS*. For each of these qualities, we have a robust and determinate conception and grasp of their simple, non-structural qualitative essences.[20] While it is certainly true that we do not conceive of these qualities via any physical-functional description which serves to fix reference, we nonetheless directly grasp them in such a way that reveals a substantive albeit *minimally descriptive* content grounded in the nature of the phenomenology itself, and which **(p.125)** we apply conceptually in introspectively characterizing those features present to our awareness, or how the property is revealed to us. This minimally descriptive content includes the intrinsic simplicity of the particular phenomenal qualities that we experience.

In being phenomenally and introspectively aware of their qualitative nature, we also may take notice of what is excluded, that is, we notice what is conspicuously absent from the phenomenology of the experience. The phenomenology lacks any trace of the heterogeneous structure or discontinuity characteristic of any candidate physical states. Take the case of the auditory experience of the musical note. If what we directly (auditorily) grasp is (as Tye contends) an external acoustic wave pattern, then the immediate object of our *q*-awareness is (roughly) a low-level physical state consisting of a pattern of vibrations of clusters of air molecules (states of compression and rarefaction) generating a complex waveform comprised of waves with varying wavelength, frequency, pressure, and intensity. But this messy, inhomogeneous, highly composite structure—which is indeed what our experience *represents*—comes to be completely glossed over in the phenomenal character or feel of the experience itself. Similar things can be said of a bluish visual experience of the ocean. If the phenomenal quality we have direct access to is in truth a wavelength reflectance disposition or profile involving water's absorption of longer wavelengths of light and its reflectance (off a watery surface) of shortwave light energy, then, contrary to what the visual phenomenology suggests, the object we phenomenally grasp is a sharply discontinuous, partite-structured relational property. In both cases, however, the phenomenal feel of our experience strongly points to directly grasped, simple, homogeneous, non-structural features that, in an ordered array, are the immediate objects and content of our awareness.[21]

The argument remains essentially the same in response to others, such as McLaughlin, who allege that the objects of *q*-awareness are internal neuro-scientific properties.[22] Returning to the example of the musical note, acoustical information from the waveform is processed by the cochlea and the auditory pathways and structures of the nervous system. This is an event involving a highly structured concatenation of electro-chemical impulses and exchanges taking place among billions of neurons in the auditory **(p.126)** cortex, along with highly complex chemical activity taking place across cell membranes and synaptic connections. Yet this auditory experience has a simple phenomenal structure, given in the musical note itself. At the phenomenal level, the event lacks any remotely recognizable isomorphic structure corresponding to the neuro-physical (let alone micro-physical) complexity. The felt sound quality itself wholly lacks any of the grainy, particulate, discontinuous spatial-temporal structure or composition of its alleged neuro-physiological and ultimately micro-physical correlates.

The point is not simply that one's access to phenomenal qualities glosses over a micro-physical structure which is somehow hidden from one's awareness, but rather that the substantive, positive grasp one does have of them rules out the possibility that the *immediate objects* of one's phenomenal awareness have such structure at all. If one wants to claim that our immediate access to phenomenal properties “smooths out” their structure to the point of completely masking their underlying nature, while at the same time granting that the phenomenological data do present us with seemingly primitive non-structural properties, then it looks like we are back again to the problem of phenomenal modes of presentation, since there is no appearance/reality distinction defined over the immediate objects of phenomenal awareness. We grasp phenomenal qualities as they really are, and we grasp them as simple, non-structural properties in the way previously described. The phenomenology of perceptual experience exhibits the positive feature of intrinsic simplicity and wholly lacks the structure and particularity that belongs to physical-functional states. The presence and absence of these features strongly argues in favor of distinguishing phenomenal character from the underlying physical-functional properties described by physics and neuroscience.

To rebut the point we are making, it is not sufficient to claim, as Tye and McLaughlin repeatedly do, that the reason we fail to directly access the underlying physical structure of phenomenal qualities is that phenomenal concepts are not physical-functional concepts, and hence we don't conceive of phenomenal qualities in that way.[23] It is more or less a truism to say *that*. **(p. 127)** Tye and McLaughlin claim that in order to be (introspectively) aware of something *as* an F, you must bring it under the concept of an F. That is, they endorse something like the following principle:

> *Introspection-as (IA)*: For subjects S, and objects or properties Q and F, if Q is revealed (in introspection) *as an F* to S, then S brings Q under the concept of F.

If phenomenal properties are revealed to us in introspection as physical-functional properties (and if we grasp them as such), then we must conceive of them via physical-functional concepts. But since, on the phenomenal concept view, we don't grasp them via physical-functional concepts, it should come as no surprise to discover that phenomenal qualities are not revealed as physical-functional properties. Now, we don't doubt that introspective awareness of and attention to phenomenal qualities requires the deployment of appropriate concepts. Certainly, when you introspect your experience of the ocean and attend to the bluish quality in your visual field, you form and apply an appropriate phenomenal concept that affords you a direct and substantive grasp of the blueness you are attending to. In so doing, you need not, and indeed do not, directly grasp the blueness *as* a physical-functional property. Nor do you directly grasp it *as* a *non*-physical-functional property. Instead, you form and apply certain corollary, rudimentary concepts to the property that are directly grounded in what it's like for you to experience it, immediately "read off," as it were, from the very content of the phenomenal concept itself. And that content derives from how the phenomenal quality is directly revealed to your awareness, i.e. as its own mode of presentation. So when you grasp phenomenal blue under a phenomenal concept B, then even though B is neither physical-functional nor *non*-physical-functional, still B is not neutral regarding the nature of the phenomenal property in question. On the basis of B, you can go on to correctly apply concepts such as "is non-structural," "is a uniform quality," and so on. That is because B itself, and the phenomenal blueness you directly grasp, has a substantive content that licenses correct application of those corollary concepts. This explanation of direct access is consistent with *IA*; that is, a non-reductivist or primitivist can grant that Q's being revealed to S as a non-structural feature requires that S exercise the concept "being a non-structural feature" in grasping it. On this view, we can introspectively grasp the nature of phenomenal qualities in part, and the remaining issue is the question of whether or not that nature is consistent with the property being a physical-functional one.

What we have been arguing is that, with regard to the issue of whether or not our special form of access to phenomenal properties is compatible with their being physical-functional properties, it is not strictly relevant that we **(p.128)** lack direct awareness of a phenomenal quality's alleged physical-functional structure under some physical-functional conceptualization; rather what matters is that the qualitative content we are positively and robustly aware of gives us a rudimentary grasp of it, which rules out any physical-functional structure as being a component of the quality we access. If the relevant physical-functional properties to be identified with phenomenal qualities are inherently structural and complex, as most theorists think they must be, then the qualities revealed to us in phenomenal and introspective awareness cannot be physical-functional features.

IV. The Argument Against a Posteriori Physicalist Accounts of Direct Phenomenal Awareness Vindicated

The purpose of the foregoing has been to sketch a plausible position midway between the extremes of demanding a full descriptive content for phenomenal properties or settling for no substantive content at all. We find the story told here to be an attractive one, cohering well with the phenomenology of perceptual experience. In fact, as we argue in this section, if we press for further clarity on (iv), insisting that we don't directly grasp the nature of phenomenal qualities, then in the end we will be forced to deny the phenomenological data. Recall that physicalists such as Tye and McLaughlin allege that:

> iv. S does not directly access Q's nature or essence.

The *nature* of phenomenal properties is not in any way revealed to us in phenomenal or introspective awareness. The phenomenal concepts we exercise in grasping phenomenal qualities have no descriptive content whatsoever, not even minimally descriptive or substantive content in the way outlined in the previous section. The *way phenomenal qualities essentially are* is not revealed to us, otherwise this would lead to affirming their intrinsic simplicity. Since they are not revealed to our awareness via any descriptive or substantive mode of presentation, the only thing it can mean to say that phenomenal qualities “are revealed as they really are” and that they “serve as their own modes of presentation” is that they are grasped by direct reference and demonstration. Tye and McLaughlin allude to this function of phenomenal concepts in several places.[24] Phenomenal concepts “do not conceptually **(p.129)** reveal anything about the essential nature of phenomenal properties; they simply name or demonstrate them.” So the bluish look of the ocean, how it phenomenally looks to you—call this phenomenal quality *b*—does not reveal anything of its nature, but is nevertheless revealed directly by your naming or demonstrating *b*, which itself is a physical-functional property. *b* is not presented to you *in* any way or *as* anything, including either a structural or non-structural property.

There are a variety of ways to think of demonstrative reference. One type of view posits two distinct states: a state C consisting of the exercise of a phenomenal concept, along with a distinct phenomenal quality or state Q that C names or demonstrates. Another view involves what we might call *self*-demonstration, whereby an instance of Q comprises the very phenomenal concept C whose deployment functions as a reflexive device, referring directly to itself. The former kind of demonstration is the sort of view Tye and McLaughlin have in mind, while the latter kind is what Block, Balog, and Papineau have in mind on their "quotational" accounts of phenomenal concepts. Each of these views can be combined with one of two broad notions of direct reference. The first is a kind of reference or demonstration that serves to reveal substantive content and information about the nature of what is demonstrated. On a theory which holds that subjects are literally *acquainted with* phenomenal properties, substantive demonstration of Q via the exercise of C is grounded in the intimate ontological and epistemic relation that obtains between the subject and the phenomenal content of her awareness. On the other notion of direct reference, demonstration does not function to reveal information about the nature of a given phenomenal property, but merely acts as a pointer or indexical for the property referred to. In this case, the demonstrating or naming function is most likely implemented through some sort of distinctive causal-functional role. The particular role of a phenomenal concept state C that directly designates or demonstrates Q will have a content along the lines of "caused by Q" or "this/that state Q." But since C contains no substantive information about Q (i.e., the subject does not access anything about the nature of Q via her exercise of C), "Q" acts as a mere label for whatever state C (rigidly) designates or demonstrates.

Given their insistence on (iv), it is clearly this latter notion of direct reference and demonstration that is in play for Tye and McLaughlin. But if that is what direct reference (and ultimately direct phenomenal access) comes down to, then their view cannot explain the phenomenally rich access we have to the qualitative character of experience. And as a result, the view of phenomenal access under consideration has no principled means of accounting for categorical differences in phenomenal character and content. Suppose **(p.130)** I look at two large colored squares painted on a wall, a blue square above a red square. (Suppose also that my color experiences are correlated in the normal way with different underlying physical properties.) I first introspectively attend to the blueness of the blue square, and then to the redness of the red square, and in doing so, I exercise phenomenal concepts that directly demonstrate the phenomenal qualities I grasp. (Call the properties B and R, and the phenomenal concepts that designate them B* and R*, respectively.) My introspective awareness of the two squares clearly seems to afford me a qualitatively different grasp of their respective colors, as my two visual experiences differ in a categorical way with respect to their phenomenal character; in one case I am phenomenally aware of B, in the other case of R. Here is the problem: when I introspectively attend to the color qualities of the squares, if there is no minimally descriptive or substantive content that I grasp via my exercise of B* and R*—no determinate conception or grasp that reflects the way the phenomenal color properties actually are, but only demonstration—then how am I to account for the qualitative differences in the way the squares look to me, that is, the differences between B and R? The differences in the way the squares look to me appear to be categorical, qualitative differences in the nature of the phenomenal properties presented. But if demonstration or naming affords no positive, determinate grasp of the way the qualities B and R are in themselves, then there exists no explanatory ground for the substantive difference in phenomenal content I notice when I deploy B* and R*.[25]

One might attempt to dispute the above point by claiming that judgments of phenomenal difference can be reliably triggered in other ways, e.g. by the appropriate physical-functional states underlying or realizing B* and R*. There is some experimental evidence indicating that, say, accurate visual discriminations need not be grounded in phenomenal discriminations.[26] Discriminating the blue square from the red square need not involve grasping **(p.131)** any categorical differences in the essences of phenomenal blue and phenomenal red. Yet our judgments about, say, color patches typically *do* seem to be grounded in this way, by noticing categorical phenomenal differences. That is, at least, often the way things phenomenally seem to us (and our argument has been broadly introspective from the start). And that observation is compatible with evidence that shows such judgments need not be, or are not always, so triggered. Materialist accounts of phenomenal concepts, like that of Tye and McLaughlin, imply that such judgments are never grounded in categorical phenomenal differences.

On the recognitional-demonstrative accounts of Tye and McLaughlin, and given their insistence on (iv), from the standpoint of the subject and the phenomenal concept being deployed, it does not matter whether it is B or some other property R that is the one being demonstrated; a phenomenal concept's having the form “this state R” instead of “this state B,” is a semantic distinction that makes no phenomenal difference in my grasp of the two properties. For it is a mere formal difference in label or naming, one that does not reflect any difference in the substantive content or nature of B and R. Intuitively, the difference between exercising a phenomenal concept that designates or demonstrates B and one that designates R is a difference in the natures of the particular phenomenal qualities demonstrated by the concepts or tokened in the concepts. That is, what marks the difference in my phenomenal grasp of B and R is the way those qualities are intrinsically, in their essence. Unfortunately, this explanation is unavailable to Tye and McLaughlin and other proponents of materialist accounts of phenomenal concepts who endorse (iv).[27] Hence we conclude that this type of account lacks the resources altogether to explain the manifest qualitative differences that exist between the what-it's-like or phenomenal feel that accompanies one's grasping type-distinct phenomenal properties. Pushing the implications of (iv) leads to implausible denial of the phenomenological evidence by failing to acknowledge the intrinsic simplicity of phenomenal character. It allows only a watered-down sense in which phenomenal properties “directly present” themselves, a very attenuated understanding of (i); and this in turn seems to imply that the character of phenomenal awareness is deeply illusory, contrary to the avowals of physicalists who aim to provide a phenomenally realistic account.

(p.132) In applying a recognitional-demonstrative account of phenomenal concepts to the problem of direct awareness, physicalists such as Tye and McLaughlin carve out a finely nuanced position that affirms and walks a thin line between the following two claims:

i. S directly accesses Q as it truly is in itself.
iv. S does not directly access Q's nature or essence.

The natures of the properties we are directly aware of are completely hidden from our awareness, not presented to us in any way at all. We have argued that physicalists cannot plausibly hold both (i) and (iv) together. (i) does justice to the phenomenal data only insofar as it allows a limited revelation thesis (R) and hence excludes (iv). Drawing out the implications of (iv) leads to an implausible denial of the phenomenological evidence and has the consequence that much of the content of phenomenal experience is illusory.

On the other hand, a *primitivist* theory of phenomenal character—one that proposes ontologically simple phenomenal qualities as the immediate objects of awareness, along with an ontologically primitive acquaintance relation to ground our access to these qualities—effectively dispels the tension that arises for physicalist theories. This is because primitivism rejects a tenet that proves to be the needless source of the tension, namely:

ii. Q is a physical-functional property.

The phenomenal feel of sensory awareness presents us with ontologically simple, non-structural qualities with a determinate intrinsic character. These are “non-physical-functional” in the sense that they cannot be characterized in terms of or reduced to basic, underlying physical properties countenanced by a mature physics. A theory is needed that incorporates this feature of the intrinsic simplicity of phenomenal character (*IS*). Primitivism effectively integrates central claims about the nature of phenomenal access—e.g., that we grasp phenomenal properties as they really are, and that we have a substantive conception of their character, albeit one that amounts to much less than a sophisticated theoretical description. Primitivism holds that in phenomenal and introspective awareness we come to access phenomenal properties as they truly are, under no distinct modes of presentation, and via a substantive (and perhaps minimally descriptive) partial grasp of their natures. On this view, to say that a phenomenal property is presented to a subject *as it genuinely is* means that the nature of the property, at least in part, is revealed through an intimate acquaintance the subject has with one or more of its instances, in a way that rules out widespread illusion with respect to the way things seem phenomenally to the subject. Such a **(p.133)** view accords well with what we find phenomenal experience to be like, with the manner in which the qualitative character of experience is immediately present to our awareness. There seems to be no other plausible explanation for what phenomenal character *is*, if it is not the essence of whatever it is that is revealed to us in *q*-awareness.

Returning to our argument from direct awareness, we discover that the very feature of intrinsic simplicity provides a basis for a direct argument for primitivism about phenomenal character. We can construct this argument by modifying the earlier failed argument of section II:

1. S directly accesses Q as it truly is in itself.
2. S directly accesses Q as a primitive, non-structural property.
3. If S directly accesses Q as it is in itself and S directly accesses Q as a property of type F, then Q is a property of type F.
4. Therefore, Q is a primitive, non-structural property.

The controversial premise—the one that will be contested by Tye, McLaughlin, and other like-minded physicalists—is (2). The basis of (2), as we have tried to argue, is simply what is given in perceptual phenomenology: we have an unmediated awareness of phenomenal character, an awareness that does not admit of an appearance/reality distinction with respect to the way things are phenomenally, and the content of that awareness is presented to us as various unanalyzable, non-structural features having a determinate qualitative character. As such, they are ontologically distinct from the neuro-physical-functional states of the brain. The result is a robust dualism of properties for which we seek an explanation.

V. The Dualist Argument From Consciousness Recast As a Fine Tuning Argument

To this point, we have been concerned to establish the central assumption underlying the argument from consciousness: the thesis that unfolding conscious experience involves systematic correlations of non-identical neuro-physical and phenomenal states. Let us call this premise the *Dual Property Correlation Thesis* (DPC). We have examined the most sophisticated recent materialist challenge to DPC and argued that it fails. We now turn to consider the argument from conscious experience as a whole.

Swinburne (2004) and Adams (1987) both argue that there can be no—or, more cautiously, that it is very unlikely there can be a—systematic scientific or natural connection between physical properties and experiential qualities **(p.134)** that would explain why they are correlated in the patterns that they are. Since an explanation of some sort is required, and the prospect for a naturalistic explanation is not in the offing, a theistic design-style explanation is the most plausible remaining option.

Note that, as stated, DPC is hardly an appropriate candidate for scientific explanation. Just restricting our attention to human experiencers, there are a huge variety of actual and possible phenomenal states, on the one hand, and neuronal state configurations in relevant portions of the brain, on the other. Present science knows a great deal about neurons, the basic elements of neuro-physiological architecture, and some of the mechanisms of communication between them. Much is still unknown about the relevant structure-types that subserve psychological operations, but no one doubts that a systematic theory is there to be had. One persuaded of the ontological irreducibility of phenomenal qualities will naturally suppose that a mature scientific classification of them would similarly discern fundamental properties and relations and specify a corresponding combinatorial algorithm. With such a classification scheme in hand, one could then hope to identify general laws of neural-phenomenal co-variation that are much simpler than brute enumeration of the billions of one-one correlates.

Swinburne dismisses the possibility of any such simplifying account on the grounds that phenomenal qualities are not quantifiable, which is a central characteristic of scientific laws.[28] However, as we have just suggested, it is plausible that qualia exhibit geometrical relations and admit variations of degree (in similarity and difference), intensity, extent (of a visual field), and so on.[29] Their essentially private nature may preclude exactitude in measurement, as a practical matter, but the argument from consciousness concerns the in-principle availability of a scientific explanation of correlations, not the practical feasibility of doing so.

Adams allows that qualia may admit such quantifiable descriptions, but believes that two insuperable obstacles to a scientific account of consciousness remain. First, the occasional flirtation with panpsychism aside, it is not plausible that any kind of phenomenal properties correlate with anything other than highly structured neural states, or at least states that bear relevant functional similarities to neural states. An hypothesized law, then, would at most chart relations between basic phenomenal qualities and certain physically complex kinds, rather than similarly basic physical kinds. As Adams sees it, such a law would merely *describe* rather than explain an important fact needing explanation, viz., that phenomenal qualities are correlated with **(p.135)** specifically neural states, and not with other physical states, whether simple or complex.[30] The second obstacle to a scientific explanation of conscious experience he expresses thus:

> …even if we had, from a purely phenomenal point of view, a single uniquely valid spectrum for each sensory modality, we would still face the mind-boggling problem of finding a mathematical relationship between the qualia of the different modalities. And without such a relationship, our law of nature will not explain why certain brain states produce phenomenal qualia such as red, yellow, and blue, and others produce qualia such as sweet, sour, and salty.[31]

We believe that Adams is correct that these two considerations preclude a certain *type* of explanation, a type of explanation that we are accustomed to seeing science pursue and often achieve, but we deny that they preclude any form of explanation whatsoever. The form of explanation that non-panpsychist qualia realism precludes is a reductive and maximally unified explanation. Our most general science, comprised of the branches of fundamental physics, has achieved great success toward the goal of a highly unified science of the elementary causal factors structuring the universe. It is an open empirical question whether a *maximally* unified theory—a single, simple equation governing a single physical constituent—might be possible for the behavior of at least relatively disorganized and isolated physical systems. Suppose a theory of this sort is not attainable. We take it that this fact alone would not cry out for a further and inevitably purposive style of explanation. If at bottom there were, say, two fundamental particles whose interactions were described by a couple of equations, then we should say that these facts provide the deepest explanations of the universe's dynamics. These facts are not themselves explainable, but even that needn't, if we are causal realists, be taken to indicate a kind of *brutely* contingent behavior. We may say that the equations encode the dispositional profiles of the particle types, profiles that are essential (de re, not merely de dicto) to those particles. Of course, it will be scientifically unexplained why the universe is ultimately constituted by these sorts of particles, rather than some other sorts, or none at all. But this fact is not germane to the present discussion, as Adams and Swinburne do not intend the argument from consciousness to be just a restricted version of the cosmological argument from contingency, à la Leibniz.[32]

Admitting primitive phenomenal qualities that causally interact with certain kinds of structured physical states is also to give up the aspiration for **(p.136)** maximal theoretical unification. It means that there are properties that are *emergent* in a robustly ontological sense, with correspondingly irreducible laws that chart their patterns of instantiation and their contributions to the dynamical evolution of physical systems. But emergence needn't be understood as an objectionably brute affair of primitive properties occurring without any causal antecedents. If we are causal realists, we shall say that fundamental properties have two distinct types of causal dispositions: the familiar “locally determinative” dispositions that are manifested in even small-scale systems (such as the disposition of negatively charged particles to repel one another, absent countervailing forces) and a disposition in the right kind of organized system to contribute to the generation of an emergent property. That fundamental physical properties confer such emergent dispositions is a surprising “hidden” fact from the standpoint of one who studies the behavior of fundamental particles in the context of isolated, small-scale systems, as particle physicists do. But it is really no more mysterious than whatever turns out to be the most fundamental *locally* determinative dispositions of matter. Both kinds of dispositions are discovered empirically by successful application of theories that posit them. Neither are transparently “intelligible” to human inquirers in the way Leibniz and Spinoza imagined.[33] If we ask why negatively charged particles are disposed to repel each other, there will be the true, even if not reductively illuminating, answer: the disposition to such interactions is the nature of the property of negative charge.[34] Likewise, if we ask why neural state N1 gives rise to an experience of phenomenal blue, rather than phenomenal yellow, there will be a true answer involving a fundamental disposition of the fundamental constituent properties of N1 toward just such an effect in just such a context, a disposition that is essential to them.

Do the foregoing considerations indicate that the argument from consciousness is altogether worthless? In our view, such a negative verdict would be too hasty. There *is* an explanandum suited to design-style hypotheses in the neighborhood, it's just that Swinburne and Adams haven't characterized it properly. We have suggested that the phenomenal realist may reasonably suppose the existence of basic, general laws connecting neural-state types and families of phenomenal-state types (corresponding more or less directly **(p.137)** to distinct sensory modalities). Such laws will encode in part facts about specific emergent dispositions of fundamental physical particulars. Here is where we see the potential for design-style reasoning. It seems plausible that there are a variety of ways things might have been with respect to the fundamental constituents of the world. We do not have in mind the Humean claim that the very particulars there are might have interacted in fundamentally different ways. We mean, rather, that there might have been ever so many different sorts of entities having different sorts of basic dispositions from the ones that are manifested in our world. And in particular, it seems a priori rather unlikely that fundamental physical entities should have emergent dispositions toward phenomenal qualities. (That this is a plausible claim is suggested by the fact that many rash but otherwise reasonable philosophers judge the emergentist view to be an utterly implausible hypothesis about our own world, and some are tempted to declare it outright impossible.) Yet, given theism, it seems more to be expected, since we may reasonably suppose the conditional probability of there being agents capable of the kind of experiential life that we enjoy on the hypothesis of theism to be at least not very low, since it is reasonable to think that one of the goods a purposive world designer would wish to see in its creation are creatures of just that sort.[35]

Taken this way, the facts underlying DPC are plausibly suited to contribute to the evidence base for the modern *fine-tuning argument* for design.[36] This argument is rooted in a vast range of information from cosmology and fundamental physics, appearing since the mid-1970s. Physicists have documented dozens of fundamental respects in which our universe, according to **(p.138)** present theory, is exquisitely “fine-tuned” for the eventual appearance of biological (and so intelligent and sentient biological) life. What is meant by the neutral term “fine-tuned” is that there are apparently contingent features of the universe, involving such things as ratios of basic particles or forces or the specific value of very large numerical constants in dynamical laws, such that had any one of them differed appreciably, the universe would not have evolved in a way consistent with the appearance of biological life. The details have been documented in several places.[37]

It is currently estimated that there are more than thirty examples of fine-tuned features of the universe, many of which appear to be independent of the others. Current physics, of course, is rife with speculation about structures more fundamental than those that current theory treats as basic. Informed opinion appears to be that some of the seemingly arbitrary constants will be explainable in more fundamental terms, but it is quite unlikely that all will be. It is noteworthy that some current speculative models which promise to explain away some fine-tuned features introduce others (as is the case with both chaotic inflationary theory, which has received significant empirical support, and the at present purely speculative string theory, which contemplates a great simplication of the basic ontology of physics).[38] We will here presume that this opinion is justified.

What, then, are we to make of all this? According to the defender of the fine-tuning argument, we should infer that our universe is fine-tuned because it was created by an intelligent, powerful designer, who created the right elements in just the right initial conditions as to allow for the eventual emergence of intelligent and sentient life. It is not improbable, and perhaps is more probable than not, that such a being would desire a universe containing some form of intelligent life. The alternative hypothesis is that the one universe there happens to be exists as a brutely contingent fact. But this is overwhelmingly *improbable*. Think of all the variations on the above constants as specifying a space of possibility. Those possibilities that are life-sustaining occupy an incredibly tiny portion of this space. That the one realized possibility happens to fall into this tiny, yet highly significant portion calls out for explanation. Assuming that the existence of such an intelligent designer is not as improbable a priori (prior to the consideration of fine tuning) as the existence of the universe itself on the chance hypothesis, the most plausible explanation is in terms of intelligent design.

(p.139) A version of the argument may be expressed in the formal apparatus of Bayesian confirmation theory as follows.[39] Let:

> L = a universe's basic parameters must be finely tuned for intelligent life to exist in it.
> E = Our universe permits intelligent life.
> D = There is a supernatural universe designer.

Where “P(A/B)” is to be read as “the probability of A, given that B is true,” the argument's proponent contends that:

> P(E/L & ~D) is *very* low.
> P(E/L & D) is significant.
> P(D/L) is significantly greater than P(E/L &~D).

That is, what we observe (E) is much more likely on the assumption that there is a supernatural designer (D) (given that L is in fact true) than on the denial of that assumption. Bayes's theorem tells us that this should lead us to upgrade the probability we assign to D (and correspondingly downgrade the probability of ~D) proportionate to the strength of the last two conditional probabilities and the weakness of the first. When a bit of evidence is highly unexpected on one theory but not so unexpected on another, the second is to be judged more likely, provided its prior probability is not so low as to swamp the significance of these conditional probability differences.

Now, we will not here try to assess the overall cogency of this argument, which has been the subject of a growing body of literature.[40] We simply suggest that the argument from consciousness is best developed as adding to the data of fine-tuning. Given the role that comparative probability judgments play in the fine-tuning argument, it is doubtful that the a priori probability of the causal properties of the universe's fundamental constituents being such as to ground physical-phenomenal correlations is sufficiently low *in itself* to significantly bolster the case for a designer. Yet, if our argument in sections III–IV above are sound, the data that consciousness provides are distinctive elements of that case insofar as they are not subject to being explained away through future development of more comprehensive physical theories. Hence, the Dual Property Correlation Thesis is best **(p.140)** taken as one important and comparatively secure ingredient in an overall cumulative case for theistic design.

References

Bibliography references:

Adams, Robert (1987), *The Virtue of Faith and Other Essays in Philosophical Theology* (New York: Oxford University Press).

—— (1994), *Leibniz: Determinist, Theist, and Idealist* (New York: Oxford University Press).

Balog, Katalin (unpublished ms.), “The ‘Quotational’ Account of Phenomenal Concepts.”

Barrow, John D. and Tippler, Frank J. (1986), *The Anthropic Cosmological Principle* (Oxford: The Clarendon Press).

Block, Ned (2006), “Max Black's Objection to Mind-Body Identity,” in Torin Alter and Sven Walter (eds.), *Phenomenal Concepts and Phenomenal Knowledge* (Oxford: Oxford University Press).

—— and Stalnaker, Robert (1999), “Conceptual Analysis, Dualism, and the Explanatory Gap,” *Philosophical Review* 108(1): 1–46.

Carter, Brandon (1974), “Large Number Coincidences and the Anthropic Principle in Cosmology,” in M. S. Longair (ed.), *Confrontation of Cosmological Theories with Observational Data* (Dordrecht: D. Reidel), 291–8.

Chalmers, David (1996), *The Conscious Mind* (Oxford: Oxford University Press).

Collins, Robin (2005), “Fine-Tuning Arguments and the Problem of the Comparison Range,” *Philosophia Christi* 7(2): 385–404.

—— (unpublished ms.), “The Many-Worlds Hypothesis as an Explanation of Cosmic Fine-Tuning: An Alternative to Design?”

Colyvan, Mark, Garfield, Jay, and Priest, Graham (2005), “Problems with the Argument from Fine-Tuning,” *Synthese* 142: 325–38.

Dretske, Fred (1995), *Naturalizing the Mind* (Cambridge, MA: The MIT Press).

Gertler, Brie (2002), “Explanatory Reduction, Conceptual Analysis, and Conceivability Arguments About the Mind,” *Noûs* 36(1): 22–49.

Gillett, Carl and Loewer, Barry (2001) (eds.), *Physicalism and its Discontents* (Cambridge: Cambridge University Press).

Horgan, Terry and Tienson, John (2001), “Deconstructing New Wave Materialism,” in Gillett and Loewer (2001), 307–18.

—— (2002), “The Intentionality of Phenomenology and the Phenomenology of Intentionality,” in D. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings* (Oxford: Oxford University Press), 520–33.

Jackson, Frank (1982), “Epiphenomenal Qualia,” *Philosophical Quarterly* 32: 127–36.

Johnston, Mark (1992), “How to Speak of Colors,” *Philosophical Studies* 68: 221–64.

(p.141) Kimble, Kevin (2006), *The Intentional Structure of Phenomenal Awareness* (doctoral dissertation, Indiana University).

Kripke, Saul (1980), *Naming and Necessity*, 2nd edn. (Cambridge, MA: Harvard University Press).

Loar, Brian (1997), “Phenomenal States,” in Ned Block, Owen Flanagan, and Guven Guzeldere (eds.), *The Nature of Consciousness: Philosophical Debates* (Cambridge, MA: MIT Press).

Locke, John (1690), *An Essay Concerning Human Understanding* (London).

Lockwood, Michael (1993), “The Grain Problem,” in Howard Robinson (ed.), *Objections to Physicalism* (Oxford: Clarendon Press, 271–91).

Manson, Neil (ed.) (2002), *God and Design* (London: Routledge).

McLaughlin, Brian (2001), “In Defense of New Wave Materialism: A Response to Horgan and Tienson,” in Gillett and Loewer (2001).

—— (2003), “Color, Consciousness, and Color Consciousness,” in Quentin Smith and Aleksandar Jokic (eds.), *Consciousness: New Philosophical Perspectives* (Oxford: Oxford University Press).

Milner, D. and Goodale, M. (1995), *The Visual Brain in Action* (New York: Oxford University Press).

O'Connor, Timothy (2000), *Persons and Causes: The Metaphysics of Free Will* (New York: Oxford University Press).

—— (2008), *Theism and Ultimate Explanation: The Necessary Shape of Contingency* (Oxford: Blackwell).

Papineau, David (2002), *Thinking About Consciousness* (Oxford: Oxford University Press).

Pruss, Alexander (2005), “Fine- and Coarse-Tuning, Normalizability, and Probabilistic Reasoning,” *Philosophia Christi* 7(2): 405–24.

Smart, “Jack” J. C. (1959), “Sensations and Brain Processes,” *Philosophical Review* 68: 141–56.

Swinburne, Richard [1979] (2004), *The Existence of God*, 2nd edn. (Oxford: Oxford University Press).

—— (1996), *Is There a God?* (New York: Oxford University Press).

Tye, Michael (1995), *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind* (Cambridge, MA: The MIT Press).

—— (2000), *Consciousness, Color, and Content* (Cambridge, MA: The MIT Press).

—— (2002), “Phenomenal Character and Color,” *Societa Italiana Filosofia Analitica, E-Symposium on Color, Consciousness, and Content* ⟨http://host.uniroma3.it/progetti/kant/field/tyesymp.htm⟩.

—— (2003), “A Theory of Phenomenal Concepts,” in A. O'Hear (ed.), *Minds and Persons* (Cambridge: Cambridge University Press).

White, Stephen L. (2010), “The Property Dualism Argument,” in Robert C. Koons and George Bealer (eds.), *The Waning of Materialism* (Oxford: Oxford University Press), 89–114.

Notes:

(1) Swinburne (1979, 2004) and Adam (1987). Their arguments develop a line of thought presented in cruder fashion by John Locke in the fourth book of *An Essay Concerning Human Understanding* (1690).

(2) For a recent challenge to the now-orthodox bifurcation, see Horgan and Tienson (2002) and Kimble (2006).

(3) See Loar (1997), McLaughlin (2001), Papineau (2002), and Balog (unpublished ms.).

(4) See Dretske (1995) and Tye (2000).

(5) Higher order varieties of representationalism typically do not attempt a reductive account of phenomenal character. Higher order theories (e.g., Rosenthal's “higher order thought,” Papineau's “higher order perception,” and Lycan's internal monitoring theories) purport to explain the awareness *relation*, why a subject is consciously aware of qualia in some cases but not in others. Such theories attempt to explain the nature of phenomenal character itself only insofar as they combine the higher order component with some other theory (e.g. first order representationalism, functionalism, etc.). We do not discuss higher order theories separately here.

(6) Cf. O'Connor (2000: 110).

(7) Key statements and discussions of arguments for property dualism are found in Smart (1959), Jackson (1982), Kripke (1980), Chalmers (1996), White (2010), Loar (1997), Block and Stalnaker (1999), and Gertler (2002).

(8) For purposes of illustration, our examples of both external world properties and internal neuro-physical properties are used somewhat interchangeably, as the points we make apply equally to both p-representational properties and type-physicalism.

(9) Loar (1997); McLaughlin (2001 and 2003); Tye (2003).

(10) The only authors who explicitly address the issue are McLaughlin and Tye, although some of what others (e.g. Loar 1997; Block 2006) have to say in the context of addressing the problem of the explanatory gap have a bearing on the issue at hand. We will focus on McLaughlin's and Tye's accounts, as well as their related comments.

(11) McLaughlin (2001: 324–6), (2003: 148).

(12) Tye (2002).

(13) “In first approximation, a phenomenal concept C refers to a phenomenal quality Q via C's being the concept that is exercised in an introspective act of awareness by person P if, and only if, under normal conditions of introspection, Q is tokened in P's current experience and because Q is tokened” (Tye 2003).

(14) Tye (2003). The examples he uses in this section primarily concern the application of phenomenal concepts to pain, but he explicitly intends his account to generalize to all phenomenal states.

(15) Block (2006) also seems to understand these ways of articulating the issue to come down to the same thing. He claims that his theory “involves a notion of a phenomenal concept that has some affinities with the ‘directness’ story in which there is no metaphysical mode of presentation at all. The doctrine that the metaphysical mode of presentation is the same as the referent says there is no metaphysical mode of presentation over and above the referent.”

(16) Horgan and Tienson (2001: 311). Although they couch their argument in the language of phenomenal concepts and conceivability, the essence of the argument could be preserved using the “awareness” locutions we have adopted.

(17) McLaughlin (2003: 147).

(18) Tye (2002, 2003); McLaughlin (2001: 324).

(19) The term “revelation” originally comes from Mark Johnston (1992), who used it to refer to the thesis that the natures of external color properties are revealed to us in our visual perception of color.

(20) The fact that many observers are not cognitively sophisticated enough to have concepts such as “simple” and “non-structural” is irrelevant to our argument. What is relevant is that so cognizing it would be an appropriate and correct judgment which is directly grounded in the character of the phenomenology. Certainly, one may grasp the “bluish” essence or character of a phenomenal quality while not recognizing it as either simple or non-structural. But in a more sophisticated exercise of one's introspective abilities, one can go on to probe further as to what kind of feature the intrinsic bluishness is. And arguably, it can be discerned as a smooth, homogeneous, non-structural feature. In our awareness of the “feels” of experience, the way things seem *is* the way things are, because there is no appearance/reality distinction defined over such features.

---

(21) Note again that we are not denying the eminently plausible claim that phenomenal qualities *represent* such candidate physical properties in the external world and that these facts constitute our being *perceptually* aware of such external objects and their properties. Our point is that they cannot be the immediate objects of *phenomenal* awareness. There is no appearance/reality distinction defined over the immediate objects and contents of *q*-awareness.

(22) Lockwood (1993: 274) makes a similar point.

(23) It is important to be clear about what is and what is not in dispute here. As proponents of non-conceptual content, Tye and McLaughlin deny that the very act of grasping a phenomenal property is concept-involving. But Tye and McLaughlin go beyond this rather innocuous claim by further denying that in grasping a phenomenal property (introspectively) via a phenomenal concept, we grasp it in a way that truly captures any part of its nature. The only concepts or judgments licensed by the phenomenal concept would be judgments that do not (non-trivially) characterize the nature of the phenomenal quality itself. (Certain trivial or relational claims about its essence could be allowed, such as “this experience is not bluish like that one,” or “this experience is identical to itself.”) Hence Tye and McLaughlin deny minimal or partial descriptivism.

(24) McLaughlin (2001: 324). See also Tye (2003).

(25) It won't work to locate the phenomenal differences in the distinct conceptual roles of B* and R*, for the original tension among (i)–(iii) re-emerges as a problem about how phenomenal character can be plausibly identified with these functional roles. I am directly and immediately aware of the qualitative character of my experience in a rudimentary yet substantive way that precludes its being merely a functional property of the sort suggested here. Moreover, conceptual roles seem by themselves *phenomenally blank* and therefore cannot constitutively ground the intrinsic, qualitative character of experience. B* and R* have distinct conceptual roles, to be sure, but the differences in qualitative character are either grounded in the differences in intrinsic natures of the properties that are their realizers (quotational account) or else grounded in the differences in natures of the properties they designate (recognitional account).

(26) See for example, the work of Milner and Goodale (1995).

(27) The problem described in this section equally confronts (materialist) “quotational” or self-referential accounts of phenomenal concepts that endorse (iv). While the functional mechanism involves self-reference or self-demonstration, this mechanism does not confer information about the nature of the properties in question; it simply functions as a sort of indexical of the form “this state I am in now (B).”

(28) Swinburne (2004: 204).

---

(29) Cf. Chalmers (1996: 224–5).

(30) Adams (1987: 255).

(31) Ibid. (257).

(32) One of us has addressed this question elsewhere. See O'Connor (2008).

(33) See Robert Adams's discussion of Lebniz's claim that an “effect is conceived through its cause” (1994: 152ff.).

(34) We needn't here enter into the dispute between pure dispositionalists, who hold that properties just are clusters of dispositions, and others who hold that dispositions are one aspect of properties alongside another, a certain qualitativity that is distinct for each distinct property.

(35) An anonymous referee objects as follows: even if we grant that the probability that there are conscious agents given theism is relatively high, there is no reason to think that conscious experience has to include primitive, non-structured phenomenal properties. So, there is no reason to think that the probability that there would be primitive, non-structured phenomenal properties given theism is significant. We make two comments in reply: (1) Granted, we have not argued above that the account of conscious experience we give is *necessarily* true of any phenomena properly termed “conscious experience.” But it seems to us plausible to think that it is necessarily true, if true at all, given that our concept of conscious experience is closely tied to general features of our own experience. It seems even more plausible that conscious experience of this sort, assuming it isn't the only sort, is especially valuable, and so something that a divine being would choose to instantiate. (2) But we needn't in any case commit ourselves to either of these claims. For the conditional probability that is at issue doesn't primarily concern the likelihood of there being primitive vs. structured higher-level properties. Rather, it is that, given that conscious properties in our world are primitive, the likelihood that such properties would arise from the lower-level structured neuro-physical base properties—i.e. that such psycho-physical correlations would hold—is higher on theism than the denial of theism.

(36) Several of the paragraphs to follow are adapted from O'Connor (2008).

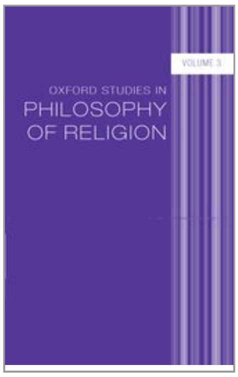
(37) Brandon Carter (1974) first called attention to this data. For a comprehensive treatment, see J. D. Barrow and F. J. Tippler (1986).

(38) Robin Collins develops this point in “The Many-Worlds Hypothesis as an Explanation of Cosmic Fine-Tuning: An Alternative to Design?,” currently unpublished.

---

(39) Here we closely follow Neil Manson's presentation in the introduction to Manson (2002: 7).

(40) In addition to the essays in Manson (2002), see Colyvan, Garfield, and Priest (2005), Collins (2005), and Pruss (2005). Finally, one of us assesses the argument in O'Connor (2008: chs. 4–5).



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

Prolegomena to Any Future Physics-Based Metaphysics

Bradley Monton (Contributor Webpage)

DOI:10.1093/acprof:oso/9780199603213.003.0008

Abstract and Keywords

Philosophers of religion sometimes appeal to physics to provide support for the hypothesis that God exists. More generally, metaphysicians sometimes appeal to physics to provide support for claims about the fundamental nature of the world. But given the current state of inquiry in physics, where the two most fundamental theories are incompatible, such arguments of physics-based metaphysics are problematic. This chapter supports this line of thought by focusing on two sorts of problematic arguments, special-relativity-based arguments against presentism and big-bang-based arguments in favor of the existence of God. The chapter is not arguing that physics-based metaphysics can't be done; the chapter is just arguing that extant examples of physics-based metaphysics are flawed. The chapter closes by considering various ways that future versions of physics-based metaphysics could potentially be successful.

Keywords: physics, metaphysics, special relativity, presentism, big bang

Philosophers of religion sometimes appeal to physics to provide support for the hypothesis that God exists. More generally, metaphysicians sometimes appeal to physics to provide support for claims about the fundamental nature of the world. But given the current state of inquiry in physics, where the two most fundamental theories are incompatible, such arguments of physics-based metaphysics are problematic. I support this line of thought by focusing on two sorts of problematic arguments, special-relativity-based arguments against presentism and Big-Bang-based arguments in favor of the existence of God. I am not arguing that physics-based metaphysics can't be done; I am just arguing that extant examples of physics-based metaphysics are flawed. I close by considering various ways that future versions of physics-based metaphysics could potentially be successful.

I. Introduction

To what extent can questions in metaphysics be answered by appealing to results in physics? Some questions in metaphysics—questions regarding necessary existence, or the nature of modality, for example—have little to do with the contingent facts about our particular universe that physics attempts to discover. However, other questions in metaphysics—questions regarding the nature of space and time, or the fundamental building-blocks of matter, for example—do potentially depend on truths learned from physics. Now, if one is a scientific anti-realist, and hence holds that the aim of science is not truth, then one will be skeptical of arguments that reach metaphysical conclusions on the basis of results from physics. But for those who are sympathetic to some version of scientific realism, physics seems to **(p.143)** be providing information that is relevant to some traditional metaphysical debates about the nature of the universe.

Indeed, metaphysicians sometimes do appeal to results from physics to provide arguments in support of some particular metaphysical view they endorse (or to provide arguments against some particular metaphysical view they reject). For example, there are relativity-theory-based arguments against presentism, claims about the Big Bang in support of the cosmological argument, claims about fine-tuned constants in support of the teleological argument, and quantum-mechanics-based arguments in favor of indeterminism and holism.

What I want to point out in this paper is that these physics-based arguments rest on shaky foundations. The arguments in question aren't necessarily wrong-headed, but sophisticated reasoning needs to be added to give the arguments the plausibility that they intuitively (yet fallaciously) appear to have. This sort of sophisticated reasoning hasn't been provided in the literature; I will start to explore the prospects for giving it below.

But why is it that physics-based metaphysics rests on shaky foundations? The answer, in brief, is that the actual physical theories that are utilized in doing metaphysics of this sort are almost certainly *false*. Our two best theories of physics, quantum theory and relativity theory, are incompatible. The evidence in favor of quantum theory suggests that relativity theory is false, and the evidence in favor of relativity theory suggests that quantum theory is false.

Here is an example of what I have in mind. Some of the evidence for quantum theory (from, for example, the two-slit experiment) suggests that a particle can be in a superposition of different positions. But in general relativity, where the curvature of spacetime is based on the distribution of matter, there is no way to have a superposition of spacetimes. Also, some of the evidence for general relativity involves experiments done with precise clocks; these experiments show that clocks in strong gravitational fields run slow compared to clocks in weak gravitational fields. But according to quantum theory, ideal clocks run at the same rate regardless of the strength of the gravitational field they are in.

There are attempts by physicists to come up with a new theory that will replace both quantum theory and relativity theory—yielding proto-theories like loop quantum gravity and string theory—but that project is very much an ongoing one, without clear results as of yet. As a result, when physics-based metaphysics is done, it's relying on false theories of physics; this is why physics-based metaphysics rests on shaky foundations.

I will elaborate on this below, in the course of looking at a couple of case studies of questionable physics-based metaphysics. But before I begin, let **(p.144)** me emphasize what I am and am not trying to argue. I do not want to say that physics-based metaphysics should not be done. Instead, I just want to point out that there are some tensions in the approach that haven't been adequately recognized, let alone dealt with. By pointing out the tensions, and by providing preliminary suggestions for how they might be resolved, I am providing a sort of prolegomena to any future physics-based metaphysics.

II. Presentism and Relativity

Below, I will take up philosophy of religion issues, when I talk about the kalam cosmological argument, and in doing so I will provide a critique of extant physics-based philosophy of religion. But it's not just in philosophy of religion that one finds problematic physics-based arguments—non-theistic metaphysics has such arguments as well. Because there are such close parallels between these two sorts of arguments, I'll lay the groundwork for the philosophy of religion discussion by starting with an example from non-theistic metaphysics.

The most well-known physics-based argument in metaphysics is perhaps Hilary Putnam's argument in “Time and Physical Geometry.” Putnam starts with what he calls the “man on the street's” view of time: “All (and only) things that exist *now* are real.” He gives an argument against this presentist thesis, appealing to physics: “I shall assume Special Relativity.” By appealing to the relativity of simultaneity in special relativity, he reaches the following conclusion:

> the problem of the reality and the determinateness of future events is now solved. Moreover, it is solved by physics and not by philosophy. We have learned that we live in a four-dimensional and not a three-dimensional world.… Indeed, I do not believe that there are any longer any *philosophical* problems about Time…[1]

It is clear from Putnam's conclusion that he takes special relativity to be *true*; otherwise, he could only conclude: *assuming special relativity*, we live in a four-dimensional world.

Putnam isn't the only person to argue against presentism by appeal to special relativity; the argument is quite a popular one in the literature on time (and presentists have given a number of arguments modifying either presentism or special relativity to restore compatibility). For example, Ted **(p.145)** Sider, in his book defending a perdurance theory of persistence, gives a number of arguments against presentism, but saves what he considers to be the best for last: “I turn finally to what is often (justifiably, I think) considered to be the fatal blow to presentism: that it is inconsistent with special relativity.”[2]

The problem with this Putnam/Sider line of reasoning is that special relativity is a *false* theory, and prima facie it's not a good idea to derive metaphysical lessons for our world on the basis of a theory that doesn't correctly describe our world. The reason special relativity is false is that it makes predictions at variance with reality. For example, according to special relativity, a clock at the base of a building will run at the same rate as a clock at the top of the building (assuming that the building is in an inertial frame of reference), but in fact the clock at the base runs slower. This fact about clocks is one piece of evidence for general relativity—according to general relativity, a clock in a stronger gravitational field runs more slowly than a clock in a weaker gravitational field.

Now, the relativity of simultaneity that holds in special relativity also holds in general relativity, so perhaps Putnam and Sider simply should have given their argument utilizing general relativity. The argument wouldn't go through in quite the same way though, because the spacetime of special relativity has more symmetry than a generic spacetime of general relativity. Specifically, there are some spacetimes of general relativity where the spacetime has sufficient structure that one can use it to pick out a preferred simultaneity relation. Nevertheless, actually specifying that there is a preferred simultaneity relation is incompatible with general relativity (on the standard interpretation of general relativity, at least). Moreover, as Gödel famously pointed out,[3] there are some spacetimes of general relativity that don't even allow a preferred simultaneity relation to be picked out—one can't foliate the spacetime into spacelike hypersurfaces. So it seems that presentism doesn't fare any better on general relativity than special relativity. Still, it would be better if anti-presentists like Putnam and Sider actually went through the details of the general-relativity-based argument.

While this would be better, it still wouldn't be good enough to conclude that presentism is false. The reason is that general relativity is most likely false. General relativity and quantum theory are incompatible, and a key project in theoretical physics is to come up with a new theory that supplants both general relativity and quantum theory. The most popular version of such a theory is string theory, but we currently don't have a full grasp of **(p.146)** what string theory even says. According to physicist Lee Smolin, there are as many as 10^{500} distinct string theories, and “we understand very little about most of these string theories. And of the small number we do understand in any detail, every single one disagrees with the present experimental data…”[4] Further, according to string theory proponent Brian Greene: “Even today, more than three decades after its initial articulation, most string practitioners believe we still don't have a comprehensive answer to the rudimentary question, What is string theory?”[5] Thus, it is difficult to reach definitive metaphysical lessons about our world on the basis of string theory. I'm not saying that it couldn't be done, but I am saying that a successful physics-based argument against presentism would have to be much more sophisticated than has so far appeared in the literature.

III. The Lesson of Special Relativity?

Here is one attempt at how such an argument might go. One could argue that, even though special relativity is false, we have learned an important lesson about the world via special relativity. Specifically, we have learned that simultaneity is relative, and hence we fully expect that any successful future theory that supplants special relativity will incorporate this result. Thus, just as presentism is incompatible with special relativity, so presentism will be incompatible with any successful future theory that supplants it, and hence presentism is false.

I want to point out some potential problems with this line of reasoning. First, there are some physicists who are working on developing theories that reject the idea that simultaneity is relative. One way of solving certain conceptual problems in the development of theories of quantum gravity is to postulate a preferred simultaneity relation. While this method is not popular, the fact that it is being pursued at least shows that one can't unequivocally assert that any theory of quantum gravity will incorporate the relativity of simultaneity.[6]

Second, there are some reasons from quantum mechanics to think that perhaps simultaneity will be shown to be non-relative. Quantum mechanics predicts,[7] and experiment confirms,[8] that there are correlations between spacelike-separated events (that is, events that are simultaneous in some inertial frame of reference) that can't be accounted for via a local **(p.147)** common-cause mechanism. Prima facie, this is incompatible with relativity. Perhaps one way of resolving the incompatibility is to give up the relativity of simultaneity.[9] While this proposal is not a popular one, it is certainly a live option, and that's enough to show that we cannot unequivocally conclude on the basis of relativity theory that simultaneity is relative.

Third, there are broader conceptual reasons to think that perhaps simultaneity will be shown to be non-relative. The history of physics contains many examples where a theory has been developed with a fundamental symmetry, but the symmetry does not hold in reality. For example, all these known or proposed symmetries are broken in nature: charge/parity symmetry, time reversal symmetry, chiral symmetry, electroweak symmetry, and supersymmetry. Perhaps Lorentz symmetry will turn out to be another symmetry of a theory that is broken in nature.[10] In fact, various experiments by reputable physicists are being done to search for violations of Lorentz invariance.[11] If Lorentz invariance is violated in the right way, then there is a preferred frame of reference, and hence simultaneity is non-relative.[12]

I want to make clear that none of the above reasons would be subscribed to by the majority of physicists. Most physicists would maintain that the relativity of simultaneity is a lesson from special relativity that will not be dropped in future physics. Could this be turned into an argument against presentism? One would presumably want to establish that the physicists have good reasons for thinking that the relativity of simultaneity will hold in future theories. One would then presumably conclude that there is good inductive reason to think that simultaneity actually is relative (and hence, presentism is false). Note that this argument, even if successful, is weaker than the arguments Putnam and Sider purport to give. Putnam says that the problem of the reality of future events is *solved by physics*—but I maintain that the most he could legitimately argue is that, based on the opinions of physicists regarding how physics is going to develop, we have good reason to think that future events are real. Sider's appeal to special relativity as a “fatal blow” to presentism is similarly overstated.

(p.148) Now, perhaps I am being uncharitable here; perhaps one could come up with a knock-down argument against presentism based on how future physics is expected to develop. But no such argument has yet been given, and hence the physics-based arguments against presentism that do exist are misleading in their apparent conclusiveness.

IV. Is The Smart Money on Science?

Sider recognizes a potential problem with his special-relativity-based argument against presentism. He writes:

> Some presentists have said: so much the worse for special relativity, at least in its Minkowskian formulation. Perhaps future empirical research will bear this out, but in cases of science versus metaphysics, historically the smart money has been on science.[13]

There are multiple problems with this line of reasoning. First of all, as I've pointed out above, empirical research has already shown that special relativity is false. But more importantly, if one looks at historical examples of conflicts between science and metaphysics, it's not at all clear that one would want to endorse the view of (then-)current science. The history of science is littered with false theories, such that one could easily reach incorrect conclusions about the world if one based one's metaphysics on these false theories.

This is an important point, so it's worth looking at a couple of examples where science-based metaphysics has gone wrong in the past. First, consider the Aristotelians versus the atomists. Aristotle's theory was the dominant theory of science, and the atomists, having little empirical evidence for their view, were arguably endorsing a metaphysical view at variance with current accepted science. We now think that the atomists' view of the universe is closer to reality. Second, consider a nineteenth-century Cartesian arguing against Newton's theory of gravity, on the basis that there is no local mechanism for gravitational influence. In the nineteenth century, the action-at-a-distance view of the gravitational force was widely accepted by physicists; the Cartesian's complaint would be viewed as a metaphysical one. But in fact, we now believe that the action-at-a-distance view of gravity is misguided, and that gravitational influence does propagate locally (specifically, at the speed of light, according to general relativity). The smart money was **(p.149)** on the Cartesian's metaphysics, not on the nineteenth-century Newtonian science.

The examples above are real-life examples of cases where one would not want to base one's metaphysics on science; the argument against science-based metaphysics can also be made by appealing to hypothetical examples. Imagine a presentist/eternalist debate during the 1800s, when Newtonian physics is the dominant theory, and suppose that the discussants, being well-versed in considerations of possibility, are aware that it's possible for the structure of spacetime to be such that there's no privileged simultaneity relation. One can imagine the presentist appealing to Newtonian physics and arguing that, since science shows us that there actually is a privileged simultaneity relation, this provides evidence that the privileged simultaneity relation corresponds to something metaphysically fundamental. While this is by no means a knock-down argument in favor of presentism, it would perhaps have some force in a time period where Newtonian physics is the dominant theory of science. But by the lights of our current science, this pro-presentism argument is misguided.

Unfortunately, Sider doesn't give any examples or argument to back up his claim that, in cases of science versus metaphysics, historically the smart money has been on science. Perhaps Sider could give such an argument, and could show that my examples where the smart money appears to be on metaphysics are just anomalies. This is one reason that my paper just provides a prolegomenon—in future arguments of physics-based metaphysics, a Sider-style appeal to the success of science versus metaphysics would have to be backed up.

V. The Big Bang and the Kalam Cosmological Argument

Now, let's turn to the second example of problematic physics-based metaphysics—an appeal to the Big Bang to support the kalam cosmological argument. The basic structure of the kalam cosmological argument is as follows:

Premise 1: The universe began to exist.

Premise 2: Everything that begins to exist has a cause of its existence.

Conclusion: Therefore the universe has a cause of its existence.

The argument goes on to establish that the cause of the universe's existence counts as God.

(p.150) What I want to focus on is the first premise of the argument. There are two lines of defense that are standardly given, a philosophical defense and an empirical defense. The philosophical defense holds that actual infinites are impossible, and hence it is impossible for a series of events to have been going on forever. To discuss that defense is beyond the scope of this paper. The empirical defense that is standardly given is physics-based. For example, William Lane Craig supports Premise 1 with the claim: “scientific evidence…indicates that the universe is finite in duration.”[14] Similarly, William Rowe asserts: “science tells us that the universe is finitely old.”[15]

What Craig and Rowe have in mind is the Big Bang hypothesis, which arises out of general relativity. The Big Bang hypothesis holds that the universe, including space and time itself, came into existence a finite amount of time ago, and shortly after the universe came into existence it was in a state of large energy density, and the energy density in the various regions of the universe has been decreasing overall. General relativity has an infinite number of models of spacetime, and in some of the models there is a Big Bang, whereas in others there isn't. Based on the empirical data we have about our universe, the models of general relativity that best describe our universe are models where there is a Big Bang.

If the Big Bang hypothesis is true, then the first premise of the kalam cosmological argument is true. This leads to the question: should we believe the Big Bang hypothesis?

In support of the Big Bang hypothesis, Craig cites a paper by four astrophysicists, with J. Richard Gott as the lead author. Gott and his co-authors write:

> the universe began from a state of infinite density about one Hubble time ago [i.e., about 15 billion years ago]. Space and time were created in that event and so was all the matter in the universe. It is not meaningful to ask what happened before the Big Bang; it is like asking what is north of the North Pole.[16]

At first glance, this passage looks like it is supporting Craig's claim. But one has to be careful here. When physicists present a theory, they may be presenting it *as true*, or they may just be presenting it as a live option, putting it on the table for consideration. According to Bas van Fraassen's understanding of science,[17] at least, physicists can *accept* a theory, and treat the theory as if it is true for the purposes of doing their science, without actually *believing* the theory. In this vein, it's worth noting that Gott and his **(p.151)** co-authors put an important caveat in their paper, a caveat that Craig doesn't quote. Gott and his co-authors write:

> That the universe began with a Big Bang is an inevitable conclusion *if* the known laws of physics are assumed to be correct and in some sense complete. It is conceivable, however, that there are laws of nature whose effects are negligible on the scale of the physics laboratory, or even on the scale of the solar system, but that might predominate in determining the behavior of the universe as a whole.[18]

So this leads to the question: should we assume that the known laws of physics are correct and complete?

The answer is: we should not. As I've mentioned, the two current fundamental theories of physics on the table, general relativity and quantum theory, are incompatible. The Big Bang hypothesis is true given the assumption that general relativity is true, but we don't know that the Big Bang hypothesis is true of the actual universe. The Big Bang hypothesis is based on a theory (general relativity) that can't accommodate evidence that supports quantum theory, and that gives us reason not to believe the Big Bang hypothesis.

In Craig's 1979 book, he doesn't seem aware of this potential problem regarding taking quantum effects into account, but by 1993 he shows more awareness of the potential problem. (In this sense, Craig is showing more sophistication than Putnam and Sider, because Putnam and Sider don't even acknowledge that developments in physics have cast doubt on the truth of the theories to which they're appealing.) Here is what Craig writes in 1993:

> During the 1980s, through the marriage of particle physics and cosmology, scientists have attempted to push back the frontiers of our knowledge of the early universe ever closer to the Big Bang.… Prior to 10^{-12} sec, however, the physics becomes speculative.… Prior to 10^{-35} sec the physics becomes extremely speculative and even unknown.[19]

I'll start with a couple of preliminary points to elucidate what Craig is talking about here, and then I'll make my main critical point.

Preliminary point #1: When Craig talks about “Prior to 10^{-12} sec,” he's talking about the time period between the Big Bang and 10^{-12} seconds after the Big Bang.

Preliminary point #2: Physicists tend to talk about stages in the development of the early universe, not in terms of the time period after the Big Bang, but in terms of the approximate amount of energy that particles in **(p.152)** the universe have at that time. So, 10^{-12} seconds corresponds to energies of 100 GeV (that is, 100 billion electron volts), whereas 10^{-35} seconds corresponds to energies of 10^{14} GeV.

Now, my main point: If one were to watch the history of the universe going backwards in time, one would see the energies increasing. Let me make the same point that Craig made about the physics getting speculative, but put in terms of energy. As the energy increases to 100 GeV, the physics becomes speculative—we're not really sure what happens at that point. As the energy increases to 10^{14} GeV (assuming it does increase to that point), the physics becomes extremely speculative, even unknown. In other words, we just don't know what happens once the energies get that high.

The way Craig puts the point, it sounds like we know that there's a Big Bang, and we know what happens in the history of the universe once 10^{-12} seconds have passed, but we don't know what happens between the Big Bang and 10^{-12} seconds after the Big Bang. But in fact our lack of knowledge is much more fundamental. Because the physics doesn't tell us what happens once we trace the history of the universe backwards in time to these high energies, we don't even know whether there's a Big Bang at all.

So given that the physics is unknown, we ought to conclude that it's unknown whether there's a Big Bang, and hence (assuming that the philosophical defense of Premise 1 is flawed) we ought to conclude that it's unknown whether the universe began to exist. An appeal to physics is not successful in showing that Premise 1 of the kalam cosmological argument is true.

VI. Objections

I'll now consider three objections to my line of reasoning in the previous section. (Similar objections could be given for my line of reasoning about presentism and special relativity, but I'll focus on the kalam cosmological argument.)

VI.1. Isn't This Scientific Anti-Realism?

Let's step back for a moment, and think about other ways one might reject the physics-based defense of the first premise of the kalam cosmological argument. If one were a strong scientific anti-realist, in such a way that one didn't think that science delivered any substantive truths about the world, then one would have clear grounds for rejecting the physics-based line of **(p.153)** reasoning—the evidence from physics would give one no reason to believe the Big Bang hypothesis. Suppose one were a more moderate scientific anti-realist, along the lines for example of van Fraassen's (1980) constructive empiricism. One would believe that the aim of science is truth about the *observable* aspects of the world. Since the beginning of the universe is presumably not an observable event, the aim of science would not include finding out whether the universe had a beginning. Note though that it would be a mistake to argue that the beginning of the universe is not an observable event simply because we couldn't exist then, because in fact we can observe events that occur before times that we could exist. For example, we can now observe the cosmic microwave background radiation, which was emitted 300,000 years after the Big Bang (assuming the Big Bang hypothesis is true). The reason the beginning of the universe is not observable is both that we couldn't exist then (given the high amount of energy in each region of the universe), and that it's physically impossible for us to observe back that far—we can't see back any further than the cosmic microwave background radiation. So a moderate scientific anti-realist like van Fraassen would also hold that the physics-based defense of the first premise is unconvincing.

What I want to make clear is that my argument from the previous section does not rely on any sympathies I might have with scientific anti-realism. According to van Fraassen's characterization of scientific realism, at least, to be a scientific realist is to believe that the aim of science is to come up with true theories, about both the observable and unobservable aspects of the world. One can be a scientific realist and still be skeptical of whether our current physical theories are true. In fact, given the current state of play in physics, even scientific realists have good reason to doubt that general relativity is true, because it does not take into account quantum effects. Because of this, it's inappropriate to appeal to general relativity as providing the theoretical framework that allows us to establish that the universe began to exist a finite amount of time ago.

To put the point dramatically: my argument only works because physics is in crisis. If there were a worked-out consistent fundamental theory of physics, scientific realists could appeal to that theory to tell them whether the universe began to exist. But since there isn't, they can't.

VI.2. Our Theories are Approximately True

Perhaps van Fraassen's characterization of scientific realism is too weak. Consider for example James Ladyman and Don Ross's characterization: **(p.154)**

> we ought to believe that our best current scientific theories are approximately true, and that their central theoretical terms successfully refer to the unobservable entities they posit.[20]

These scientific realists could say: maybe general relativity in all its details is false, but nevertheless we should expect general relativity to be *approximately* true. The realists could thus say that we should believe what general relativity tells us about the Big Bang. The realists could maintain that general relativity gets the details of what happens between the Big Bang and 10^{-12} seconds after the Big Bang wrong, but the overall picture is approximately right—the universe did start a finite amount of time ago.

The notion of approximate truth is notoriously slippery—the question that always has to be asked is: in what respects, and to what degree, is the theory approximately true? Suppose that general relativity is right in suggesting that spacetime is curved, and that the curvature of spacetime is correlated with the distribution of matter in spacetime. This would be enough to make it reasonable to consider the theory to be approximately true. This could be the case even if general relativity makes incorrect predictions for what happens at large energy scales. As one traces the history of the universe backwards in time, toward larger and larger energy density, it could be that at sufficiently large energy density the universe undergoes a bounce, and energy density starts decreasing again. This scenario is compatible with the universe having been in existence forever—the cycle of the universe expanding and then contracting could have been going endlessly. If that were the case, general relativity could reasonably be considered to be approximately true, even though the Big Bang hypothesis is false.

This endless crunch/bounce model is not just pure philosophical speculation; at least some prominent physicists consider such a scenario to be a live option. Specifically, Paul Steinhardt and Neil Turok have proposed the *cyclic model*, where the expansion phase that the universe is in now will be followed by a contraction phase.[21] At the end of the contraction phase the universe will undergo a “bounce,” and will begin expanding again. Steinhardt and Turok still use the terminology of the “Big Bang,” but for them the Big Bang is just the event of the bounce, where the universe starts expanding again. The cyclic model is compatible with the universe having been in existence forever.

For those who are familiar with the discredited oscillatory models of the 1920s, it's worth making clear that the cyclic model is not one of those models, but is instead based on up-to-date physics. It is arguably compatible **(p.155)** with M-theory, and utilizes the branes (the multi-dimensional analogue to strings) that M-theory endorses. According to the cyclic model, there are ten dimensions of space, but six of those dimensions are compactified. We are living in one three-dimensional brane, and there is one other three-dimensional brane, which is at some distance from ours as measured in the fourth uncompactified spatial dimension. The branes oscillate back and forth relative to each other, and the Big Bang corresponds to the moment when these two branes get so close that they repel each other and bounce. There are many more details to the cyclic model, and it is detailed enough that it makes empirically verified quantitative predictions, but there's no need to go into those details here. I've said enough to show that there is a respected theory under consideration by physicists which does not entail that the universe came into existence with a Big Bang.

One could legitimately hold that, according to the cyclic model, general relativity is approximately true. Indeed, many aspects of general relativity carry over to the cyclic model. But if the cyclic model is true, and the universe has been in existence forever, general relativity is not approximately true with respect to general relativity's suggestion that the universe came into existence at the Big Bang. This shows that general relativity can be approximately true, even if the Big Bang hypothesis is false. Thus, one can't appeal to the approximate truth of general relativity to establish that the universe came into existence a finite amount of time ago.

VI.3. We Should Work With the Best Theories We Have

The third and final objection I want to consider goes as follows. I haven't seen it in print before, but I have heard it in discussion, so here is my best representation of the objection:

> Where else should we look for our guidance for our metaphysics, if not our best fundamental physical theory? Unless and until general relativity is replaced by a better theory, we should just assume that general relativity is true, for the purposes of doing metaphysics. It would be unreasonable to ask scientists and engineers to send rockets into space using a non-existent physical theory; the best they can do is to use the best fundamental physical theories we have. The same sort of reasoning holds for the metaphysician. Thus, in evaluating the first premise of the kalam cosmological argument, we should assume that the best theory we have regarding the beginning of the universe is true.

This is perhaps the line of thought that Mauro Dorato has in mind when he gives a physics-based argument against presentism. He rejects special-relativity-based arguments because special relativity is false, and instead gives a **(p.156)** general-relativity-based argument. He recognizes that general relativity may well be false as well, but in a footnote he says “until a reasonably agreed upon quantum theory of gravity is available, we can assume that [general relativity] *is* a fundamental physical theory.”[22] Dorato doesn't go on to say *why* this assumption is appropriate, but perhaps I've represented above what he has in mind.

I'll point out two problems with this line of reasoning. The first problem is that, in the absence of some sort of hedging, someone reasoning this way could end up having contradictory beliefs. General relativity is not our only fundamental physical theory—quantum theory is also standardly considered by physicists to be a fundamental physical theory. Thus, if we're going to assume, for the purposes of evaluating the kalam cosmological argument, that general relativity is true, we should also assume that quantum theory is true. But general relativity entails that the structure of spacetime is correlated with the distribution of matter, whereas quantum theory has a fixed spacetime. If one assumes that both general relativity and quantum theory are true, one will end up believing that the structure of spacetime both is and is not correlated with the distribution of matter, and hence have contradictory beliefs. It would be arbitrary to assume that general relativity is true and not quantum theory, because they are standardly considered to be equally fundamental theories. (Or at least, one would need a worked-out argument to explain why it's legitimate to assume that general relativity is true but not quantum theory, and that's an argument proponents of the kalam cosmological argument haven't given.) But if one assumes both are true, one ends up believing a contradiction.

Even setting this logical problem aside, there is another reason one wouldn't want to simply assume that general relativity is true for the purposes of doing metaphysics. The reason is that metaphysics isn't meant to be an elucidation of our best current scientific theories; metaphysics is meant to get at truth. In order to rely on our current best scientific theories in doing metaphysics, one would need to argue that these theories are true. As we've seen, given the state of current physics, it is unreasonable to simply assume that general relativity is true.

Moreover, if philosophers assume that general relativity is true, they won't be able to make sense of much of the work that's being done in contemporary theoretical physics. Why would physicists spend so much effort to develop string theory, if general relativity is true? Philosophers who took this approach would be viewed by physicists as naive.

(p.157) But perhaps my focus on truth is misguided—perhaps it is wrong to construe metaphysics as a search for truth. Perhaps metaphysics is simply out to elucidate the fundamental structure of the world, under the supposition that our best fundamental physical theories are true. This can be thought of as counterfactual metaphysics: what metaphysical claims would be true, were general relativity true? Assuming that there are no problems with the rest of the kalam cosmological argument, the argument can establish that it would be true that God exists, were general relativity true. This would be an interesting and important philosophical result, but we have to recognize what the limitations of the result are. It would be of the same importance as establishing that God would exist, were Cartesian physics true, or that God would exist, were Ptolemaic astronomy true. While these are philosophically interesting results, they don't get at reality—or at least, one would need an independent argument to show that they are getting at reality, even though the physical theory in question is false. Proponents of the kalam cosmological argument, historically at least, have tried to establish that God actually exists, not that God would exist if it were the case that some particular scientific theory is true.

VII. Bringing in Probabilities

The kalam cosmological argument is presented as a deductive argument, and so far I have been treating it as such. If we aren't warranted in believing that all the premises are true, then we aren't warranted in believing that the argument is sound. But what happens if we treat it instead as an inductive argument—one that leads to an increase in the probability assigned to the hypothesis that God exists?

Imagine an agent who fully believes that the kalam cosmological argument is valid, and fully believes that the second premise is true, but doesn't know whether the first premise is true—this agent isn't sure whether the universe began to exist. Also, suppose that this agent starts out not knowing any physics. Now, suppose that the agent becomes an expert in physics—her beliefs about physics match the beliefs that top physicists have. Will this lead to an increase in the probability she assigns to the truth of the first premise, and hence, will this lead to an increase in the probability she assigns to the hypothesis that God exists?

My answer to these questions is: it depends. But seeing why that's the right answer will help us better to understand the inductive version of the kalam cosmological argument.

(p.158) Let's first focus on the conclusion of the argument, the proposition that God exists. It could be that, before learning any physics, the agent already fully believes that God exists; she already assigns probability 1 to that proposition. As a result, learning the physics is not going to lead to an increase in the probability for that hypothesis.

Let's suppose then that, before learning any physics, the agent assigns a non-extremal probability, between 0 and 1, to the proposition that God exists—call that proposition G. For similar reasons, let's assume that the agent assigns a non-extremal probability to the hypothesis that the universe began to exist—call that proposition B. Note that the probability that the agent assigns to G, $P(G)$, is greater than or equal to the probability that the agent assigns to B, $P(B)$, since the agent fully believes that B, along with other propositions the agent fully believes, entails G.

Now, let E be the set of beliefs that expert physicists have. The key question we want to ask is: when the agent learns E, does this lead to an increase in the probability she assigns to G? In other words, is it the case that $P(G|E) \rangle P(G)$?

The answer is: it depends on what agent we're discussing. I'll describe two sorts of agents, each of whom is rational and has intuitively reasonable probability assignments, but where for one agent $P(G|E) \rangle P(G)$, whereas for the other agent $P(G|E) \langle P(G)$.

First, imagine an agent, Alice, who is an atheist—she initially assigns a low probability to G. Alice starts out not knowing any physics, and she thinks it highly improbable that the universe began to exist. As a result, she initially does not find the kalam cosmological argument at all plausible, even though she believes the second premise, and believes that the argument is valid. Moreover, suppose that the kalam cosmological argument provides the only reason she has to believe that the universe has a cause of its existence. But then, she learns physics, and learns that the hypothesis that the universe began to exist is a live option, an option taken seriously by expert physicists. This leads her to increase her probability assignment to the hypothesis that the universe began to exist, and hence to increase her probability assignment to the hypothesis that the universe has a cause of its existence. She believes that this cause would be God, and hence learning the physics E leads her to increase her probability assignment to G.

Now, imagine an agent, Bob, who is a believer—he initially assigns a high probability to G. Bob starts out not knowing any physics, and he thinks it highly probable that the universe began to exist. Moreover, the kalam cosmological argument provides the only reason that Bob believes in God. You can see how the story goes: when Bob learns physics, he learns that the hypothesis that the universe has been in existence forever is a live option, an **(p.159)** option taken seriously by expert physicists. This leads him to decrease his probability for B, and hence to decrease his probability for G.

The lesson I draw is that the evidence from physics can, in conjunction with the kalam cosmological argument, lead to an increase in the probability that a rational and reasonable agent assigns to *G*; to this extent the inductive version of the kalam cosmological argument is effective. However, the same evidence and argument can lead to a decrease in the probability assigned to *G* for a different rational and reasonable agent; to this extent the argument is counterproductive (from a theistic perspective).

Is that all there is to say? Isn't there an objectively right answer as to whether the inductive kalam cosmological argument is successful? I have been implicitly utilizing a subjectivist interpretation of probability in this section, and while I think that is the right interpretation of probability to endorse, this is a controversial matter. Could a different interpretation of probability yield a more definitive answer as to whether the inductive kalam cosmological argument is successful?

I'm not going to try to answer these questions here—these are the sorts of questions one would have to address if one wanted to give a successful physics-based inductive version of the kalam cosmological argument. Perhaps there is an objectively right way to fill in the blank in the following statement: given the current state of physics, the probability that the universe began to exist is ___. But without seeing the argument, I am doubtful.

VIII. Other Arguments

The problems with physics-based metaphysics are not restricted to the special-relativity-based argument against presentism and the big-bang-based argument in support of the existence of God. In the interest of concision, I won't go into details regarding other problematic arguments; I'll just give a couple of examples so the reader sees what I have in mind.

I'll start with another non-theistic example of physics-based metaphysics, and then I'll give another example from philosophy of religion. Sometimes one encounters vague claims that physics has pretty much gotten things right, even though it has not yet found a potentially true most fundamental theory. For example, Frank Jackson writes:

> it is reasonable to suppose that physical science, despite its known inadequacies, has advanced sufficiently for us to be confident of the *kinds* of properties and relations that are needed to give a complete account of non-sentient reality. They will be **(p.160)** broadly of a kind with those that appear in current physical science, or at least they will be as far as the explanation of macroscopic phenomena go…[23]

Jackson doesn't say anything to defend these claims, and it's not at all obvious to me that these claims are true. Why is Jackson so confident that the theory that supplants general relativity and quantum theory won't have new kinds of properties and relations that are used in explaining some macroscopic phenomenon? I simply don't see any reason to think that. While that may be true, I'd want to see the argument, and as far as I know no such even semi-compelling argument has been given.

Here is another example from philosophy of religion of problematic physics-based metaphysics. Peter van Inwagen, in the context of discussing the problem of evil, writes:

> there is at least good reason to think that a deterministic world that contains complex life—or any life at all—may not be possible. Life depends on chemistry, and chemistry depends on atoms, and atoms depend on quantum mechanics…and, according to the “Copenhagen interpretation,” which is the interpretation of quantum mechanics most favored by working physicists, quantum mechanics is essentially indeterministic.[24]

There are at least three problems here. First, while it is part of the content of the Copenhagen interpretation that quantum mechanics is indeterministic, it's not part of the content of the Copenhagen interpretation that quantum mechanics is *essentially* indeterministic. That is, it's not part of the content of the Copenhagen interpretation that there are no deterministic interpretations of quantum mechanics that are empirically equivalent to the Copenhagen interpretation. Or at least, if that is a part of the content of the Copenhagen interpretation, then the Copenhagen interpretation is demonstrably false, since David Bohm's theory is a deterministic interpretation of quantum mechanics that is empirically equivalent to the Copenhagen interpretation.[25]

The second problem with what van Inwagen says is that he seems to be assuming quantum mechanics is true when he says “atoms depend on quantum mechanics.” In fact quantum mechanics is false, since it makes predictions at variance with the evidence standardly taken to support general relativity. So either atoms exist in the actual world, in which case they don't depend on quantum mechanics, or atoms don't exist, in which case chemistry doesn't depend on atoms.

(p.161) The third problem with what van Inwagen says is that he's illegitimately assuming that complex life could only exist in a world with our sort of physics when he says that there's good reason to think that it's *impossible* for there to be a deterministic world with complex life. Even if our world is in fact indeterministic, it doesn't at all follow that it's only possible for complex life to exist in a world with our sort of physics. I conclude that van Inwagen has presented another example of problematic physics-based metaphysics.[26]

IX. Looking Forward

My criticism is of extant arguments of physics-based metaphysics. This leads to the question: how could metaphysicians of the future successfully appeal to physics? By way of closing, I'll summarize seven potential options they have for doing so:

1. One option is to simply wait for progress in physics to be made. If a prima facie successful theory of quantum gravity is developed, then physics will no longer be in the position of having two most fundamental conflicting theories. At this point, the debate over physics-based metaphysics will more closely parallel the debate over scientific realism generally.
2. Another option for metaphysicians to successfully appeal to physics is to make explicit that their arguments are based on opinions about how physics will go. There is some prima facie force to the idea that, if almost all physicists think that future theories of physics will make certain claims about the world, then it is reasonable for philosophers to conclude that those claims about the world are true. But the argument would need to be more developed, and so far, metaphysicians who appeal to physics have not developed that argument.
3. A third option is for metaphysicians to restrict the scope of metaphysics. They can say that metaphysics isn't about discerning fundamental truths about the world, but instead is about establishing conditional claims, of the form: if that physical theory were true, then the world would be such-and-such way. Even setting aside issues of physics- **(p.162)** based metaphysics, there is some evidence that metaphysicians reason this way. For example, consider David Lewis's line that set theory “offers an improvement in what Quine calls ideology, paid for in the coin of ontology.”[27] One can read this as part of the project of different metaphysicians offering competing packages regarding the nature of the world, without providing any truth-conducive grounds to favor one over the other. Obviously this is a controversial reading of metaphysics (one that Lewis himself would probably not accept), but it is one in which simplistic physics-based arguments could find a home.

4. Another option for metaphysicians to successfully appeal to physics is to limit the scope of the conclusion they're drawing from physics. For example, in the relativity-based arguments against presentism, the metaphysicians could refrain from asserting that relativity theory is true, but could instead just assert that it's *possible*. They could then argue against those presentists who think that presentism is necessarily true. This is an interesting line of argument, but it's quite different from the arguments that, for example, Putnam and Sider give. Putnam and Sider, recall, are trying to show that presentism is false in the actual world, not that there's some possible world where presentism is false. Moreover, this new line of argument is itself controversial. Presentists could argue that special relativity is in fact impossible; or they could simply concede that eternalism is possibly true; or they could hold that presentism is true of any world with time, and that a world where special relativity is true is timeless.

5. A fifth option for metaphysicians to successfully appeal to physics is to give an argument analogous to one that Kurt Gödel gives.[28] Gödel argues for the ideality of time (and, implicitly, against presentism) with the following line of reasoning. He points out that there are some spacetime models of general relativity where there is no passage of time, and suggests that those who hold that there is passage of time in other spacetime models of general relativity are committed to holding that "whether or not an objective lapse of time exists (i.e. whether or not a time in the ordinary sense of the word exists), depends on the particular way in which matter and its motion are arranged in the world."[29] Gödel then says: "This is not a straightforward contradiction; nevertheless, a philosophical view leading to such consequences can hardly be considered as satisfactory."

Whether Gödel's argument is sound or not is a subject of much controversy. All I want to point out here is that metaphysicians **(p.163)** could potentially give an analogous argument to justify their particular argument of physics-based metaphysics. For example, for the relativity-based argument against presentism, the metaphysician could hold that there is a lot of evidence for *something like* special relativity, and could hold that whether presentism is true shouldn't depend on the details of whether special relativity is true, or whether a theory closely related to special relativity is true.

6. The sixth option for metaphysicians to give successful arguments of physics-based metaphysics is to appeal only to uncontroversial aspects of physics. This is arguably what Guillermo Gonzalez and Jay Richards have done in their pro-intelligent design argument.[30] In brief, their argument holds that there is a correlation between habitability and observability—the regions of the universe that are best suited for life are also the regions of the universe that are best suited to make scientific observations. They conclude that this provides evidence for the existence of a designer.

Setting aside the question of whether their argument is successful, what is interesting to note is that they aren't appealing to controversial physics to give their argument. Instead of making the strong claim that some particular fundamental physical theory is true, they are simply appealing to more empirically oriented facts about where life could flourish in the universe, and what sorts of observations could be made in different regions of the universe. This wouldn't help with the kalam cosmological argument, but it does provide a different way to give a physics-based argument for the existence of God.

7. The final option I'll consider for how to do physics-based metaphysics is for metaphysicians to appeal to some form of structural realism. They can recognize that our extant fundamental theories of physics are false, and yet can hold that such theories surely get something about the structure of the world right. This move is forthrightly made by Ned Hall:

> I think that we in fact possess detailed knowledge of the structure and internal nature of molecules and atoms. At the same time, I have no confidence whatsoever that our best current physics is anywhere close to right about the fundamental nature of the physical world. I conclude that, somehow or other, our empirically well-grounded knowledge of molecular and atomic structure must be the sort of thing that can survive wholesale revisions in our conception of the fundamental natures of the constituents of molecules and atoms. For example, we can be perfectly and justifiably confident that methane **(p.164)** has a tetrahedral structure, even if our current best explanation of what it is for it to have that structure is fatally flawed. It's an interesting puzzle how exactly we can manage to be in such an epistemic situation.[31]

> It's not obvious to me that we are in such an epistemic situation, but if we are, I agree that it's an interesting puzzle regarding how we can manage to be in that situation. If we can discern aspects of the fundamental structure of the world from false physical theories, then some physics-based metaphysics can be done. From an epistemological standpoint, it would be nice to know how such discernment is possible. (Hall doesn't attempt to provide an answer, and I haven't seen anyone else successfully do so.)

So, to sum up: I am not saying that physics-based metaphysics can't be done. What I am saying is that metaphysicians who want to appeal to physics successfully are going to have to be more sophisticated than they have been in giving such arguments.

References

Bibliography references:

Aspect, Alain, Grangier, Philippe, and Roger, Gérard (1982), “Experimental Realization of Einstein-Podolsky-Rosen-Bohm *Gedankenexperiment*: A New Violation of Bell's Inequalities,” *Physical Review Letters* 49: 91–4.

Bell, John Stewart (1964), “On the Einstein-Podolsky-Rosen Paradox,” *Physics* 1: 195–200.

Bohm, David (1952), “A Suggested Interpretation of the Quantum Theory in Terms of ‘Hidden Variables’ I,” *Physical Review* 85: 166–79.

Craig, William Lane (1979), *The Kalam Cosmological Argument* (London: The MacMillan Press).

Craig, William Lane (1993), “The Finitude of the Past and the Existence of God,” in William Lane Craig and Quentin Smith (eds.), *Theism, Atheism, and Big Bang Cosmology* (Oxford: Oxford University Press).

Dorato, Mauro (2002), “On Becoming, Cosmic Time and Rotating Universes,” in Craig Callender (ed.), *Time, Reality & Experience* (Cambridge: Cambridge University Press), 253–76.

Gödel, Kurt (1949), “A Remark about the Relationship between Relativity Theory and Idealistic Philosophy,” in P. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist* (New York: Tudor), 557–62.

(p.165) Gonzalez, Guillermo and Richards, Jay (2004), *The Privileged Planet: How Our Place in the Cosmos is Designed for Discovery* (Washington, DC: Regnery Publishing).

Gott, J. Richard, Gunn, James, Schramm, David, and Tinsley, Beatrice (1976), “Will the Universe Expand Forever?,” *Scientific American*, March: 62–5.

Greene, Brian (2004), *The Fabric of the Cosmos: Space, Time, and the Texture of Reality* (New York: Alfred A. Knopf).

Hall, Ned (2007), Review of Wesley Salmon's *Reality and Rationality, Notre Dame Philosophical Reviews*, ⟨http://ndpr.nd.edu/review.cfm?id=8524⟩, archived at ⟨http://www.webcitation.org/5Ytn0VK0B⟩.

Jackson, Frank (1998), *From Metaphysics to Ethics: A Defense of Conceptual Analysis* (Oxford: Oxford University Press).

Kostelecky, V. Alan (2004), “Gravity, Lorentz Violation, and the Standard Model,” *Physical Review D*, 69: 105009; preprint available at ⟨http://arxiv.org/pdf/hep-th/0312310v2⟩.

Ladyman, James and Ross, Don (2007), *Every Thing Must Go: Metaphysics Naturalized* (Oxford: Oxford University Press).

Lewis, David (1986), *On the Plurality of Worlds* (Oxford: Blackwell).

Maudlin, Tim (1996), “Space-Time in the Quantum World,” in J. Cushing, A. Fine, and S. Goldstein (eds.), *Bohmian Mechanics and Quantum Theory: An Appraisal* (Dordrecht: Kluwer), 285–307.

Maudlin, Tim (2007), “Why Be Humean?,” in *The Metaphysics Within Physics* (Oxford: Oxford University Press).

Monton, Bradley (2006), “Presentism and Quantum Gravity,” in Dennis Dieks (ed.), *The Ontology of Spacetime* (Dordrecht: Kluwer), 263–80.

Pospelov, Maxim and Romalis, Michael (2004), “Lorentz Invariance on Trial,” *Physics Today*, July: 40–6.

Putnam, Hilary (1967), “Time and Physical Geometry,” *Journal of Philosophy* 64: 240–7.

Rowe, William (2005), “Cosmological Arguments,” in William Mann (ed.), *The Blackwell Guide to the Philosophy of Religion* (Oxford: Blackwell Publishing), 103–16.

Sider, Theodore (2001), *Four-Dimensionalism: An Ontology of Persistence and Time* (Oxford: Oxford University Press).

Smolin, Lee (2006), The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next, (Boston, MA: Houghton Mifflin).

Stanwix, Paul, Tobar, Michael, Wolf, Peter, Locke, Clayton, and Ivanov, Eugene (2006), “Improved Test of Lorentz Invariance in Electrodynamics Using Rotating Cryogenic Sapphire Oscillators,” *Physical Review D* 74: 081101.

Steinhardt, Paul, and Turok, Neil (2002), “A Cyclic Model of the Universe,” *Science* 296: 1436–9.

Tooley, Michael (1997), *Time, Tense, and Causation* (Oxford: Oxford University Press).

Van Fraassen, Bas (1980), *The Scientific Image* (Oxford: Oxford University Press).

Van Inwagen, Peter (2006), *The Problem of Evil* (Oxford: Oxford University Press).

Notes:

(1) Putnam (1967: 247).

(2) Sider (2001: 42).

(3) Gödel (1949: 557–62).

(4) Smolin (2006: xiv).

(5) Greene (2004: 376).

(6) For more detail on this point, see Monton (2006).

(7) Bell (1964).

(8) Aspect et al. (1982).

(9) For more discussion of this move, see for example Maudlin (1996: 295–8) and Tooley (1997: 358–62).

(10) See for example Pospelov and Romalis (2004: 40) for an endorsement of this line of thought.

(11) For one representative experiment, see Stanwix et al. (2006); for a list of over forty such experiments, see ⟨http://www.physics.indiana.edu/~kostelec/faq.html⟩, archived at ⟨http://www.webcitation.org/5YtluOslD⟩.

(12) See Kostelecky (2004: 3) for some discussion of the circumstances in which a violation of Lorentz invariance would imply a preferred frame of reference.

(13) Sider (2001: 42).

(14) Craig (1979: 140).

(15) Rowe (2005: 113).

(16) Gott et al. (1976: 65).

(17) Van Fraassen (1980).

(18) Gott et al. (1976: 65).

(19) Craig (1993: 67–9).

(20) Ladyman and Ross (2007: 68).

(21) Steinhardt and Turok (2002).

(22) Dorato (2002: 255).

(23) Jackson (1998: 7).

(24) Van Inwagen (2006: 118).

(25) Bohm (1952).

(26) If I were to continue the list, I'd endorse Tim Maudlin's (2007: 62–3) critique of David Lewis (1986): Lewis inappropriately appeals to classical physics to support Humean supervenience. I'd also endorse Ladyman and Ross's (2007: 22–7) critique of metaphysicians who inappropriately assume that fundamental physics describes the world in terms of “microbanging”—little things coming into contact with each other.

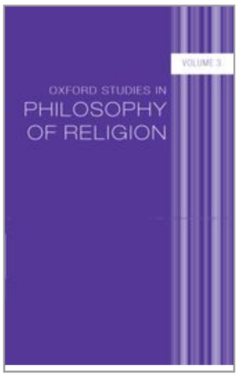
(27) Lewis (1986: 4).

(28) Gödel (1949).

(29) Ibid. (562).

(30) Gonzalez and Richards (2004).

(31) Hall (2007).



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

O'Connor's Cosmological Argument

Graham Oppy

DOI:10.1093/acprof:oso/9780199603213.003.0009

Abstract and Keywords

In *Theism and Explanation*, Tim O'Connor provides ‘an argument for the existence of a transcendent necessary being as the source and basis of the ultimate explanation of contingent beings and their interconnected histories’. This chapter argues that O'Connor's argument is unsuccessful: each of the three most plausible naturalistic views concerning ‘the ultimate explanation of contingent beings and their interconnected histories’ is more theoretically virtuous than any account that is committed to transcendent necessary beings. Thus, the chapter argues that naturalists can reasonably refuse to choose between ‘infinite regress’, ‘brute contingency’, and ‘immanent necessity’ (at least insofar as we are concerned with the ultimate explanation of contingent beings and their interconnected histories).

Keywords: contingency, cosmological argument, naturalism, O'Connor, theism, ultimate explanation

In Chapter 3 of his recent book *Theism and Ultimate Explanation*, Tim O'Connor discusses what he calls “the existence stage” of his cosmological argument.[1] According to the dust-jacket notes of the book, “O'Connor defends a novel version of the Leibnizian cosmological argument from contingency for the existence of a transcendent necessary being as the source and basis of the ultimate explanation of contingent beings and their interconnected histories.” Taking the dust-jacket notes at face value, then, we can reasonably infer that Chapter 3—which is entitled “Ultimate Explanation and Necessary Being: The Existence Stage of the Cosmological Argument”—is where we find the heavy lifting in O'Connor's cosmological argument: this is where we are provided with his argument for the *existence* of “a transcendent necessary being as the source and basis of the ultimate explanation of contingent beings and their interconnected histories.” In this paper, I aim to provide a critical assessment of the material that O'Connor presents in this key third chapter.

I. O'Connor's Project

O'Connor has rather different-sounding things to say about what he aims to do, and what he succeeds in doing, in the chapter under examination. On the one hand, in the Preface, he repeats the claim that is made on the dust-jacket of the book (though perhaps the direction of causation is the other way around!): “I defend a version of the ‘Leibnizian’ cosmological argument from contingency for the existence of transcendent necessary being as the source and basis for ultimate explanation, of contingent beings and their interconnected histories.”[2] Moreover, O'Connor goes on to outline the direction of argumentation in Chapter 3 in the following terms: **(p.167)**

> The stage is thus set for a consideration of that most fundamental metaphysical question, or set of questions, concerning existence itself. The best form of the question—one that presumes the least—is this: are there contingently existing objects, and if there are, why do those particular contingently existing objects exist and undergo the events they do? I consider a variety of options for providing an outline of an answer to this question. I argue that *the only one that is not beset by fundamental problems* is one that accepts the existence of contingent beings and maintains that they are the causal product of a purposive, transcendent necessary being, one for whom existence and essence are inseparable, one that must simply be.[3] (my italics)

On this account, it seems that O'Connor must suppose that the achievement of Chapter 3 is very considerable indeed: what Chapter 3 shows is that *there is only one* “outline of an answer” to “the best form of the most fundamental metaphysical question concerning existence” that is “not beset by fundamental problems”—namely, the “outline of an answer” provided by one kind of traditional theism.

On the other hand, at the end of Chapter 4—at the conclusion of “the identification stage” of his cosmological argument—O'Connor writes:

> I conclude that the reflections of the past two chapters provide a significant reason to believe that contingent being has its explanatory ground in necessary being…. Philosophical arguments on fundamental metaphysical theses rarely if ever settle the matter, and the present line of argument is no exception. I do contend, however, that the considerations adduced are weighty enough to warrant some serious conceptual development by the philosopher inclined towards a rival metaphysical vision.[4]

And, in the coda to the book, O'Connor says:

> It has been a long and at times wild ride. Though I promised at the outset not to arbitrarily dismiss our cab when it arrived at an agreeable location, no doubt even the most patient of my readers will agree that it is best to put it into park for a while. I have argued that the Existence Question can sensibly be posed, and if it is, we will see reason to posit necessary being at the heart of our metaphysics…. But unlike both John Duns Scotus and Samuel Clarke, the two most direct influences on the present line of argument, I would take pains once again to deflate the quasi-mathematical rhetoric of “demonstration.” It is, I believe, a plausible line of thought that adds to the attractiveness of traditional theistic metaphysics. As such, it merits careful response by those contemporary philosophers (who name is legion) who take themselves to have seen through its illusory promise—a much more careful response, alas, than is fashionable in these days of breezily dogmatic adherence to philosophical naturalism.[5]

(p.168) On this alternative account, it seems that O'Connor steps back from the bold claim that his argument shows that, for example, naturalistic responses to “the best form of the most fundamental metaphysical question concerning existence” are “beset by fundamental problems.” Rather, his more modest claim is that he has adumbrated “a plausible line of thought” that “merits a careful response” from those who are not disposed to accept the claims of “traditional theistic metaphysics.”

I propose to take O'Connor at his later word. I'm happy to pay careful attention to his “plausible line of thought.” However, I think that, when we do pay careful attention to his “plausible line of thought,” we see that we don't need to add anything “weighty” to extant dissenting treatments of this topic. Far from calling for “serious conceptual development,” O'Connor's discussion merely calls for attention to refinements of points that have already been made—again and again—by, for example, naturalistically inclined philosophers.

II. Existence Questions

O'Connor begins his discussion by offering a characterization of the central aim of metaphysics, namely “to articulate a theoretical framework that makes possible *ultimate* explanation of reality—that is, a natural or non-arbitrary stopping point…to the nested series of available plausible explanations for increasingly general aspects of the world.” Moreover, O'Connor goes on to observe that “the realisation of this aim has been thought to require an answer to the most basic of metaphysical questions: *Why is there anything contingent at all?*” Finally, O'Connor settles on a refined version of this question that he takes to be, in some sense, “best”: *“Are there contingently existing objects, and if there are, why do those particular contingently existing objects exist and undergo the events they do?”*[6] I think that O'Connor's characterization of the central aim of metaphysics is not well-suited to the investigation that he proposes. I take it that the central aim of metaphysics is to provide the most general theory of what there is and what it is like. Furthermore, I take it that when we assess the merits of rival metaphysical theories, we consider the usual range of theoretical desiderata: simplicity, scope, fit with empirical data, explanatory power, fit with well-established scientific theories, and so forth. It may be that, when we consider metaphysical theories, we think that one *prima facie* desirable feature of such **(p.169)** theories is that they should provide “ultimate explanations of reality”—but we should also recognize it may be that, in the best metaphysical theories, failure to provide “ultimate explanations of reality” is outweighed by performance on other theoretical desiderata. To insist that the central aim of metaphysics is to construct theories that provide “ultimate explanations of reality” is to make a question-begging assumption about the characteristics of best metaphysical theories (perhaps as a result of giving unwarranted primacy to the “explanatory completeness” of metaphysical theories).

It is worth noting that, in framing the discussion in the previous paragraph, I followed O'Connor in supposing that, if a theory says that the existence of the universe is a brute, unexplainable fact, then that theory provides no answer to the existence question.[7] One might well think, to the contrary, that if a theory says that the existence of the universe is a “brute, unexplainable fact,” then that theory provides one kind of answer to the existence question—and, moreover, if one followed out this line of thought, one would naturally suppose that no suitably developed metaphysical theory *could* fail to provide an answer to O'Connor's most basic of metaphysical questions. However, if we were to think about the matter in this way, then we would be required to reformulate O'Connor's understanding of the central aim of metaphysics: for, on his view, the central aim of metaphysics is to arrive at a metaphysical view on which there are no “brute, unexplainable contingencies.” No matter how much higher theories might score on grounds of simplicity, scope, fit with empirical data, fit with well-established scientific theory, and so forth, O'Connor will not allow them entry to the metaphysical contest if they involve “brute, unexplainable contingencies.”

Suppose we do accept that metaphysical theorizing should be oriented primarily by a fundamental question about existence. We are still left with the question about how exactly to formulate this question. O'Connor himself notes that “the question admits distinct formulations of greater precision”: “*Why is there anything (contingent) at all? What explains the fact that there are contingent things? What explains the fact that these contingent things exist? What explains the fact that these contingent things exist rather than those others that were possible?*”[8] Moreover, given his claim that “the best form of the question—one that presumes the least—is this: are there contingently existing objects, and if there are, why do those particular contingent objects exist and undergo the events they do?”,[9] it is clear that O'Connor would **(p.170)** grant that each of the formulations in the preceding list also admits of a “better” variant in which it is not assumed in the question that there are contingently existing objects. But, how, exactly, do we arrive at the conclusion that the best form of the question is the one that O'Connor identifies? While there is a point at which O'Connor says that he “will in due course consider the relevance and appropriateness of certain of these formulations,”[10] I do not think that he provides an explicit answer to this question.

I take it, in fact, that what O'Connor does is to endorse the most explicit form of the question that he takes it to be possible to answer. While one might think that one ought to endorse the contrastive question—*What explains the fact that these contingent things exist rather than those others that were possible?*—O'Connor is persuaded that *that* question does not admit of an answer: it cannot be that there are explanations for all contrastive facts. By contrast, O'Connor holds that it is possible to answer the correspondingly detailed non-contrastive question—*What explains the fact that these contingent things exist? Why do those particular contingently existing objects exist and undergo the events they do?*—and that is why he claims that this is the best version of the existence question.

It is, I think, worth noting that, because O'Connor justifies the formulation of his existence question in this way, he may not be entitled to ask the following rhetorical question: “Given that our natural, intuitive assent is towards our world's being ultimately intelligible (as the pervasive tendency to raise the Existence Question reveals), why should we retreat when we see where it leads us? For, if the pervasive tendency is to follow Jonathan Edwards in asking the contrastive question, then the recognition that *that* question does not admit of an answer surely does give us *some* reason to rethink any intuitive enthusiasm that we might have for the idea that the world is ‘ultimately intelligible.”[11]

III. The Range of Options

If we were to list “schematic” metaphysical frameworks, we would have to include at least the following. (In the specification of these frameworks, I use the word “the universe” to refer to the entirety of physical reality, or, at any rate, to the largest part of physical reality that includes the earth, and **(p.171)** that is in no way externally related—causally, spatially, temporally, spatiotemporally, etcetera—to any other part of physical reality. This usage takes no stance on the question whether physical reality is exhausted by what cosmologists typically refer to as “our universe.”)

1. The universe is all that exists, the universe has an initial state, and the existence and nature of the initial state of the universe is metaphysically necessary.
2. The universe is all that exists, but the universe has no initial state, and the “history” of the universe does not have a finite measure. Moreover, each state of the universe can be explained in terms of earlier states of the universe.
3. The universe is all that exists, the universe has an initial state, and the existence and nature of the initial state of the universe has no explanation.
4. There is a necessarily existent agent who intentionally brings about the existence of the universe.
5. The universe exists because a particular kind of non-agential principle obtains (e.g., as John Leslie has it, the universe exists because it should exist).

At least *inter alia*, O'Connor discusses all five of these schematic metaphysical frameworks; and, of course, he argues that the fourth should be preferred to the others. Since I'm not particularly interested in considering the relative merits of the fourth and fifth “schemes,” I shall simply ignore the criticisms that O'Connor makes of the invocation of non-agential principles in explanations of the existence of the physical universe. What I shall focus on is the case that O'Connor makes on behalf of the claim that the fourth scheme should be preferred to each of the first three schemes. In particular, I'm interested in the reasons that O'Connor gives for being dissatisfied with the first three schemes, and the reasons that he gives for thinking that the fourth “scheme” might be developed into a satisfying metaphysics. Before we turn to our examination of these schemes, it is worth noting that each of the first three plausibly scores rather better than the fourth on grounds of simplicity (ontological economy). For, not only does the fourth scheme involve an additional kind of entity, it also involves the instantiation of entirely new kinds of properties. If the fourth is to be preferred to any of the first three schemes on explanatory grounds, then it will not be merely because it provides explanation where the others fail to do so; rather, it will be because it is worth paying the price of the extra ontology and ideology in order to purchase the explanations in question.

(p.172) IV. The Initial State of The Universe as Necessarily Existent

The first comprehensive scheme that we consider supposes that the universe has an initial state whose existence and features are necessary—in the same sense in which O'Connor supposes that his intentional creator is necessary, a sense that may well be “opaque to our cognitive capacities,”[12] so that we need not suppose that we understand how it can be that the existence and features of an entity are necessary in this sense.

O'Connor does not really discuss this proposal. What he does, instead, is to criticize the suggestion that the universe is, itself, a necessary existent. Against this suggestion, O'Connor objects that nothing with mereological complexity and diversity can be necessarily existent in the sense at issue: nothing composed of elementary particles and fields could be necessarily existent “of itself.”[13] While O'Connor, himself, admits that his discussion is “a rather hard-scrabble excursion,”[14] I don't propose to discuss it in any detail; for the purposes of the present section, I can happily concede that O'Connor may be right to hold that nothing that is composed of elementary particles and fields can be necessarily existent in the relevant sense.

Clearly, then, the proposal *here* under consideration is that the universe has an initial state that has no mereological complexity or diversity: it is not composed of elementary particles and fields, and it has no topological parts. Subsequent states of the universe evolve from this initial state, under the twin influences of law and chance. Moreover, subsequent states of the universe do have mereological complexity and diversity; and at least some subsequent parts of the universe are composed of elementary particles and fields.

Further details of this account depend upon how extensive the universe turns out to be. If the initial state of the universe is something like “the initial state of the Big Bang,” then the account requires that our Big Bang emerged from an initial state that had no mereological complexity or diversity, and that was not composed of elementary fields. (Of course, on standard accounts of the Big Bang, particles don't emerge until much later on.) We needn't suppose that this initial state has to be what we get if we project a standard general relativistic model back to an initial singularity, i.e. that the initial state is identified with a point on the boundary of a standard general relativistic model. Rather, all we need to suppose is that the correct—perhaps quantum-gravitational—model of the universe is equipped with an initial **(p.173)** state that lacks mereological complexity and diversity, etcetera. If, on the other hand, the initial state of the universe is something that precedes “the initial state of the Big Bang,” then we don't yet have any bounds on our conception of what that state might have been like, or what might lie between it and our Big Bang.

Of course, it is speculative to conjecture that the universe had an initial state that had no mereological complexity or diversity; it could hardly be said that we have compelling reason to believe that this is, in fact, the case. However, the important question here is whether this theory should be preferred to a theory that postulates a necessarily existent agent that brings about the existence of the universe. Both theories are committed to the existence of the universe. One theory proposes that the initial state of the universe is necessarily existent, and hence that the initial state of the universe has no mereological complexity or diversity. The other theory proposes that there is a necessarily existent agent who brings the universe into existence, and, hence, that there is an agent with no mereological complexity or diversity. Since simplicity favours the first hypothesis, and since it seems pretty clear that one *can* reasonably hold that there is nothing that favours the hypothesis that there are agents with no mereological complexity or diversity above the hypothesis that the universe had an initial state that had no mereological complexity or diversity, it seems that the naturalist is on pretty solid ground in thinking that our first scheme should be preferred to the fourth. (Note I don't say here one *could not* reasonably hold that the hypothesis, namely there are agents with no mereological complexity or diversity, is sufficiently to be preferred to the hypothesis, namely the universe had an initial state devoid of mereological complexity or diversity, that there remains justification for the postulation of the additional ontology and ideology required by the theist. However, one would surely need to advert to other considerations, beyond those that are invoked in the context of cosmological arguments.)

V. A Beginningless Sequence of Causes

The second comprehensive scheme that we consider supposes that, if the universe is co-operative, then we can have an answer to O'Connor's existence question without postulating *any* new ontology or ideology. According to this scheme, even if the universe is all that exists, provided that the universe has no initial state, and provided that the “history” of the universe does not have a finite measure, it remains the case that we *can* explain why particular contingently existing objects exist and undergo the events they **(p.174)** do, since each state of the universe can be explained in terms of earlier states of the universe.

Against this Humean proposal, O'Connor objects:

> The crucial claim here, it seems to me, is that this form of explanation can be *complete*, leaving nothing further to be explained, *even when the explanandum is a single event of short duration*. This claim is simply mistaken. [Thus my response is different from that of Rowe…who allows that contingent events within the universe are fully explainable, in principle, by citing their natural causes, while contending that the universe as a whole may still require additional explanation.] A complete explanation would be *unconditional*—it would not appeal to factors that are themselves left unexplained. This requirement evidently is not met for local, sequential explanations where one event is explained in terms of another which itself is an unexplained given in terms of the explanation at hand…. Explanations cannot be unconditional if the terms are themselves all contingent.[15]

I'm not confident that I can reconstruct the argument O'Connor has in mind here. What he *seems* to be saying is that the Humean thinks we can give a complete explanation of an event E1 by invoking another event E2, even though we acknowledge that, for the purposes of the explanation at hand, the event E2 is just an unexplained given. (Why did the first domino fall? Because the second domino fell against the first! Why did the second domino fall against the first? Enough with the questions! I've already fully explained why the first domino fell!)

If this is what O'Connor is saying, then it doesn't square with my understanding of the Humean view. I don't think that the Humean supposes that we can give a *complete* explanation of an event E1 by invoking another event E2, even though we acknowledge that, for the purposes of the explanation at hand, the event E2 is just an unexplained given. Rather, I think what the Humean holds is that, if we suppose there is an infinite regress of events in which each is causally dependent on the next, then we are entitled to think that *there is* a complete explanation of why particular contingently existing objects exist and undergo the events they do that adverts only to the elements of the infinite regress of events. Of course, the Humean doesn't suppose that we can *give* the complete explanation—*per impossible*, that would require us to enumerate a beginningless infinite list of premises; rather, the Humean supposes we can recognize that, were it the case that there is an infinite regress of events in which each is causally dependent on the next, we would have no reason to think that our scheme is explanatorily incomplete.

(p.175) O'Connor adverts to an example from Pruss,[16] which he takes to "illustrate the essential explanatory incompleteness of simply noting the stepwise dependence within a beginningless sequence of events":

> Suppose a cannon is fired at time t_0 and the cannonball lands at t_1. Now consider the infinite sequence of momentary events spanning all times between the two events excluding t_0 and including t_1. There is no first event in this sequence, as there is no first temporal instant after t_0. Thus, though the entire sequence has a finite duration, it still meets Hume's envisaged scenario of a beginningless infinite sequence of events, each causally dependent on events that precede it. Hume should conclude that this series is explanatorily complete, but this is evidently false: the entire sequence of events has a partial explanation in terms of the firing of the cannon at t_0.[17]

I don't believe that this argument establishes the conclusion that O'Connor (following Pruss) takes it to establish. On the one hand, it is clear that Hume himself would have denied that there are infinite sequences of finite duration, since he believed in minimum spatial and temporal intervals. On the other hand, if Hume had not believed in minimum spatial and temporal intervals, then he would have amended his claim: either he would have claimed that it is only beginningless infinite sequences of events where the sequence has infinite temporal measure that give complete explanations, or he would have claimed that it is only beginningless infinite sequences of events for which it is true that there isn't anything temporally anterior to all of the members of the sequence that give complete explanations. Moreover, not only would Hume have amended his claim, but he would have offered justification for the amendment. As Pruss's example makes vivid, an infinite sequence of events, in which the sequence of events is of finite duration, can be preceded by other events (and those other events might be invoked to explain the infinite sequence of events). However, it is *impossible* for an infinite sequence of events, in which the sequence of events is of infinite duration, to be preceded by other events; and it is *impossible* for an infinite sequence of events, for which there is nothing that is temporally anterior to all of the members of the sequence, to be preceded by other events.

Suppose it is the case that there is a beginningless infinite sequence of events, where the sequence of events has infinite temporal measure; i.e. suppose that each state of the universe can be “conditionally” accounted for in terms of prior states of the universe, and that the universe is infinitely old. Compare two schemes, one saying that the universe is all that there is, and the other that adds a necessarily existent creative agent to the story. As before, **(p.176)** the addition of the necessarily existent creative agent clearly has ontological and ideological costs. But what does it buy us? In this case, we have already assumed that *each* state of the universe can be “conditionally” accounted for in terms of prior states of the universe, and that the universe is infinitely old. So where is the activity of the necessarily existent creative agent supposed to gain purchase? If there were an initial state of the universe, we could understand how there could be a role for a creative agent: the creative agent could make the initial state of the universe (and the laws). But if there is no initial state of the universe (and if the laws have *always* been around), then it seems that there isn't anything for a creative agent to do.

If the line of thought at the end of the previous paragraph is right, then it seems, in this case, on the assumption that there is a beginningless infinite sequence of events, where the sequence of events has infinite temporal measure, the naturalist is on pretty solid ground in preferring the naturalistic scheme to the theistic scheme. Again, I don't say one could not hold that, all things considered, the theistic scheme should be preferred to the naturalistic scheme; but, as before, one would surely need to advert to other considerations, beyond those that are invoked in the context of cosmological arguments.

Perhaps, it might be thought, the argument that there would be nothing for a creative agent to do went a bit fast. Many theists suppose that God has a further role “sustaining” the universe in existence, or the like. Maybe, then, what the postulation of a necessarily existent creative agent purchases is an explanation of the continuing existence of the world: it is only against the background of the necessarily existent creative agent's conserving power that an event E1 can cause an event E2.

I see at least the following two difficulties with this proposal. First, it seems that the theistic scheme now compares even less favourably with its naturalistic comparator: in order to get an alleged explanatory advantage, we postulate not only the necessarily existent creative agent, but also the need for this agent to exercise conserving power to preserve the universe in existence. Second, it is hard to see that this theistic scheme affords any advantage when it comes to the explanation of individual events: supposing that a necessarily existent creative agent exercises sustaining power could only “complete” the explanation of E2 in terms of E1 by making the necessarily existent creative agent causally responsible for E2—a result that most theists will likely find unacceptable, since it leads to problems concerning the libertarian freedom of human agents, and the like. While I agree that the argument I have sketched is incomplete, it seems plausible to me to think that the provision of further detail would not undermine the conclusion that the naturalist is on pretty solid ground in preferring **(p.177)** the naturalistic scheme to the theistic scheme, on the assumption that there is a beginningless infinite sequence of events.

VI. Brute Contingency

The third comprehensive scheme that we consider supposes the universe is all that exists, that it has an initial state, and that the existence and nature of the initial state of the universe has no explanation. As we have already noted, O'Connor is not prepared to admit this scheme to the metaphysical contest: in his view, to say that, ultimately, the particular contingent objects there are exist and undergo the events they do as the result of brute contingency, is simply to fail to answer “the existence question.” That's not to say O'Connor offers nothing by way of justification of this stance. For instance, he writes:

> *Prima facie*, it seems unreasonable to say that the question is a perfectly coherent one, and that it correctly presupposes that the universe and everything therein need not have existed—that is to say, its existence is entirely contingent—and nonetheless hold that there is no answer to it: hold that the universe's existence is simply a brute unexplain*able* fact. It seems even more unreasonable, then, to deny that, other things being equal, given two metaphysics such that one of them provides a final, non-arbitrary answer to the existence question and one of which leaves it unanswerable, we should prefer the one that answers it on account of its greater explanatory power.[18]

There are at least two obvious points to make here. First, it is clearly true that, *other things being equal*, given two metaphysics such that one of them provides a final, non-arbitrary answer to the existence question and one of which leaves it unanswerable—or even merely unanswered—we should prefer the one that answers the existence question on account of its greater explanatory power. But —as we have already taken pains to note—it is obviously not the case that “all other things are equal” when we compare the theistic scheme that O'Connor endorses with naturalistic schemes that leave the existence question unanswered or unanswerable. O'Connor's theistic scheme has ontological and ideological investments that are surplus to the commitments of the rival naturalistic schemes. So the question must be addressed whether there is sufficient justification for this additional investment in O'Connor's theistic scheme; and, at the very least, it is not *obvious* that the investment is justified simply because it purchases an answer to—or, at any rate, the possibility of an answer to—the existence question.

(p.178) Second, it is quite unclear why O'Connor thinks that it is “*prima facie* unreasonable” to hold that the existence of the universe is contingent, and yet to hold that the existence of the universe is brutely contingent, i.e. not something that admits of further explanation. When we compare this third scheme with O'Connor's theistic scheme, it seems to me that there is *prima facie* reason—and not merely *prima facie* reason—to prefer the third scheme. After all, even on O'Connor's own account, the theistic scheme has various problematic ontological and ideological commitments that do not figure in the third scheme; whence, there is surely *prima facie* reason to suppose that friends of the third scheme *can* reasonably hold that the costs of those commitments outweigh the explanatory benefits that the commitments confer.

One point that friends of the third scheme might think to make in the present context is that there is a sense in which brute contingency is unavoidable: it is necessarily the case that, if there is contingency, then there is unexplained contingency. As O'Connor notes, if we suppose that, for each truth P, there is an explanation of why P rather than Q, for any possible alternative Q to P, then it follows that there are no contingent truths.[19] Hence, if there are contingent truths, then there is at least one truth P for which there is no explanation of why P rather than Q, for some possible alternative Q to P. So, the friend of the third scheme might be inclined to say, if any “scheme” that allows for contingency is bound to contain some unexplained contingency, surely there is no serious objection to the third scheme simply on the grounds that it allows that the existence of the universe is an unexplained contingency.

In response to this kind of argument, O'Connor says:

> One could still press the Existence Question in a principled manner by urging that one should seek explanation for every fact other than those for which there is an explanation of why there can be no explanation of those facts. If an event is an outcome of a non-deterministic causal system, there will be no explanation of why it occurred rather than any of the other possible outcomes of the system (at least in many cases). But there is an explanation for why we cannot explain such contrastive facts that involves the character of the system in question. By contrast, the philosopher who maintains that the existence of the universe is a brute, inexplicable fact cannot, by the very nature of his position, explain why such a fact is unexplainable. (Even if the universe has no explanation, it could have had one, as there could have been an equally contingent being that caused it to be.)[20]

Suppose that the existence of the universe has no explanation: it is just a brute contingent fact that the universe exists. Must we suppose that even **(p.179)** though the existence of the universe has no explanation, it could have had an explanation? In particular, must we suppose that, even though the existence of the universe has no explanation, there could have been a contingent being that caused the universe to come into existence? I don't think so. Suppose, for example, that we hold a strong version of the thesis that origins are necessary: if something has a certain kind of origin, then it is not possible that that thing should have had a different kind of origin.[21] While these kinds of theses are clearly controversial, it would be a very hard saying to hold that one could not be *reasonable* in adopting them! Given this strong thesis of the necessity of origin, if it is true that *our* universe has no explanation for its existence, then it is not possible that *our* universe has an explanation for its existence. So—it seems—the friend of the third scheme can provide an explanation of the inexplicability of the existence of the universe, given only the datum that the existence of the universe has no explanation; at the very least, the friend of the third scheme has a principled reason for denying the claim that there could have been a contingent being that brought the universe into existence, given there wasn't anything that brought the universe into existence.

Suppose, the foregoing considerations notwithstanding, we agree that the philosopher who maintains that the existence of the universe is a brute, inexplicable fact cannot, by the very nature of his position, explain why such a fact is unexplainable. Should we suppose that this represents a serious problem for friends of the third scheme? I don't think so. For, despite O'Connor's implicit claim to the contrary, it seems pretty clear that the friends of his fourth (theistic) scheme are no better placed to meet the demand that a scheme should provide an explanation for every fact other than those for which there is an explanation of why there can be no explanation of those facts. (We shall fill out this argument at the end of the next section, after we have examined the details of O'Connor's fourth scheme.)

Given the arguments that have been developed in this section, it seems to me to be plausible to conclude that naturalists are on pretty solid ground in concluding that the third scheme should be preferred to the fourth scheme. Once again, I don't say that one could not hold that, all things considered, the fourth (theistic) scheme should be preferred to the third (naturalistic) one; but, as before, one would surely need to advert to other considerations, beyond those that are invoked in the context of cosmological arguments.

(p.180) One final observation. It is an interesting feature of O'Connor's position that he defends commitment to what one might call “brute necessities,” i.e., commitment to internal and external relations of necessity that are opaque to our cognitive capacities. In the foregoing discussion, I have simply followed O'Connor's lead in allowing that a commitment of this kind can be reasonable. However, it might occur to some friends of the third scheme to claim that, at the very least, there needs to be some accounting of the relative costs of commitments to “brute necessities” and “brute contingencies.” Even if we are prepared to accept that causation involves a kind of necessity that is not fully intelligible to us, we might reasonably baulk at the idea that it is better to postulate the necessary existence of a being whose necessary existence is not fully intelligible to us than it is to suppose that there is no explanation for the existence of the universe. Or, at least, so friends of the third scheme might be tempted to suppose.

VII. A Necessarily Existing Agent

The fourth scheme—the scheme of which O'Connor does say that it is “the only one that is not beset by fundamental problems”—holds that the universe is the intentional contingent creation of a necessarily existing agent. On O'Connor's own account, his treatment is “sketchy,” owing to the presence of “complicating issues”;[22] indeed, O'Connor says of the underlying theory of intentional agency that he claims “only to show that there appears to be a way in which this might go, the details of which must be developed elsewhere.”[23] Here is the core of his account of “the agency of a transcendent necessary being”:

> A personal necessary being's activity in generating a contingent order is to be thought of, in the first instance, as the direct causing of an internal state of intention that a particular determinate state of affairs obtain. This is not, importantly, to be treated as an elliptical expression for there being some prior state of the agent that brings about, in mechanistic fashion, the agent's coming to have the intention. Rather, the intention is irreducibly a product of the agent qua agent…. When we apply this generic picture to the creative activity of a transcendent necessary being, the explanatory framework it will engender is roughly this: Let it be supposed that our necessary being has some purpose whose content is P and recognises that creating contingent order C would satisfy P. Suppose further that he generates an intention whose content is *that C obtain* in order to attain P, and that C's obtaining is itself an **(p.181)** immediate product of that intention. In such a circumstance…the core activity and its product are perfectly well explained by reference to the agent's purpose and his belief that C would satisfy it.[24]

On O'Connor's account, then, we are to suppose that we have a *complete* explanation of the existence of the universe if we can advert to the necessary existence of an agent who has some purpose, who recognizes that creating the universe would satisfy that purpose, and who generates an intention whose content is *that the universe exist* in order to attain that purpose.

If the agent were only to possess the purpose contingently, then it seems that we would not have a complete explanation of the existence of the universe; similarly, if the agent were only to contingently recognize that creating the universe would satisfy the purpose, then we would not have a complete explanation of the existence of the universe. Hence, it seems we must be supposing that the agent has the purpose of necessity, and it recogniszs that creating the universe would satisfy the purpose of necessity. However, in order to make room for contingency, it seems clear we are then to suppose that the agent also has other purposes of necessity, and that it recognizes a variety of ways of satisfying each of these purposes of necessity. (Something like the former assumption is needed if we are to suppose that the agent might not have chosen to create anything; and something like the latter assumption is needed if we are to suppose that the agent might have chosen to make something other than our universe.) But, if that's right, then the picture that O'Connor is recommending is one on which there is brute contingency in the generation of the intention whose content is *that the universe obtain* in order to satisfy a held purpose: the agent has the same purposes and the same beliefs about how it could realize those purposes in all possible worlds, and yet it forms different intentions for realizing purposes in different possible worlds, even though there is *nothing else* that varies across those possible worlds that bears on the formation of those intentions.

Should we be happy with the claim that there can be brute contingency in a complete explanation? That surely depends upon exactly how we are to understand the expression “complete explanation.” O'Connor does not give an explicit account of “complete explanation.” However, there are places where he makes use of the expression: for example, he tells us that “a complete explanation would be unconditional—it would not appeal to factors that are themselves left unexplained.”[25] While it is tempting to read this **(p.182)** as an endorsement of the claim that there cannot be brute contingency in complete explanation, I think that it is better read as endorsement of the claim that there are cases in which we can properly think that brute contingencies are not unexplained.

Consider what O'Connor has to say about “indeterministic mechanistic causal processes in the natural world”:

> Tertiary syphilis, when untreated by penicillin, leads to paresis, a form of motor paralysis, about 28% of the time. Furthermore, paresis only arises in syphilitics. Assuming for the sake of example that this statistical fact is a result of a fundamentally ‘chancy’, indeterministic process, we can explain a given occurrence of paresis, nonetheless, by pointing to the presence of the untreated syphilis. (And this despite the fact that the explanandum was not even made likely by the explaining factor cited.) For the latter, *ex hypothesi*, caused the former in an indeterministic fashion. It may well be that in such a case we cannot explain why the paresis occurred rather than not, given that it was causally possible (and indeed likely) that it not occur under just those circumstances. But why is it not enough for a non-contrastive explanation of what actually occurred that we point to the causal mechanism (however chancy) that produced it?[26]

While I agree with O'Connor that there is something right in the idea that we could explain occurrences of paresis by pointing to the causal mechanism that produced it, it seems to me that it would be very odd to say that we can give a *complete* explanation in this way. What we would more naturally say is that a complete explanation has to advert both to the causal mechanism that produced the paresis, and to the fact that that causal mechanism is chancy. Why did Fred get paresis? Well, there's a 28 per cent chance that tertiary syphilis will lead to paresis, and Fred happens to be a tertiary syphilitic who lucked out. End of story. But if you tell me only that Fred has tertiary syphilis, then you plainly haven't given a complete explanation of his paresis.

I think that there is a pattern here that is repeated wherever we think that there is objective chance. In one sense, we think that objectively chancy events lack complete explanations, because those events are partly the products of objective chance. (“There's no explanation of why that particular radioactive atom decayed exactly when it did.”) In another sense, we think that objectively chancy events can be given complete explanations, but then we insist that such explanations have to make explicit mention of the fact that there is objective chance involved. (“That particular radioactive atom decayed when it did because radioactive decay is an objectively chancy process, and that radioactive atom decayed when it did simply as a matter **(p.183)** of objective chance.”) So long as we are clear about the sense in which we are using the expression “complete explanation,” no confusion results.

In the light of the foregoing discussion, we can see that O'Connor's favoured account of intentional action gives the generation of intentions metaphysical standing rather similar to the metaphysical standing of radioactive decay. Moreover, I think, the fact that we cannot make any more precise claim than that O'Connor's favoured account of intentional action gives the generation of intentions—metaphysical standing *rather similar to* the metaphysical standing of radioactive decay points to a difficulty in O'Connor's overall position. The reason why we cannot make a more precise claim about the comparison between the metaphysical standing of the generation of intentions and the metaphysical standing of radioactive decay is that it is not clear whether, on O'Connor's favoured account, it would be appropriate to say that the generation of intentions is “objectively chancy”. In particular, the most detailed statement that O'Conner makes—“we suppose that a necessary being might have sustained any of a very wide range of possible orders answering to competing motivations, none of which came decisively weighted”[27]—is silent on the question whether we should suppose there is a probability distribution over the intentions that it is possible for the postulated necessarily existent agent to form.

If there is a probability distribution over the intentions that it is possible for the postulated necessarily existent agent to form, then it is clear that O'Connor is wrong to claim that a complete explanation of the action of the agent need only advert to the generation of the intention that was actually formed: in the only sense of “complete explanation” that could properly apply in this case, the explanation would be “complete” only if it also adverted to the existence of the probability distribution. Moreover, when we compare the fourth scheme thus understood with, say, the third scheme, we have to decide whether the cost of postulation of a necessarily existent agent whose intentions are formed as a matter of objective chance is worth the alleged improvement obtained in the answer to the question *Why do those particular contingently existing objects exist and undergo the events they do?* Friends of the third scheme might well reasonably ask: why is it better to postulate objectively chancy intentions in a necessarily existent agent than it is to postulate that the existence of the universe is brutely contingent? At the very least, it is far from obvious that this is the best way of trading off the virtues of theories in the case of the existence of the universe.

On the other hand, if there is no objective probability distribution over the intentions that it is possible for the postulated necessarily existent agent **(p. 184)** to form, then it seems to me that we lose any justification for the claim that these intentions are the *products* of the agent. If A is produced by B, or if A is under the control of B, then either there is some kind of counterfactual dependence of A on B, or else there is some kind of objective probability distribution that links the production of A to B. In particular then, there is no such thing as “the irreducible production of an intention by an agent *qua* agent” in circumstances in which there is neither counterfactual dependence nor objective probability distribution to tie the production of the intention to the agent. If *that* is what is meant to be invoked in O'Connor's fourth scheme, then friends of the third scheme can rest easy: whatever problems may be involved in supposing that the existence of the universe is brutely contingent pale into insignificance in comparison with the problems that arise on the supposition that there is “irreducible production of an intention by an agent *qua* agent” in circumstances in which there is neither counterfactual dependence nor objective probability distribution to tie the production of the intention to the agent.

Before I turn to examination of some loose ends, it is probably appropriate for me to summarize the argument of this section. I have claimed that O'Connor's fourth scheme is crucially underspecified: it fails to tell us whether it is to be supposed that there is an objective probability distribution over the intentions that it is possible for the postulated necessarily existent agent to form. If it is not to be supposed that there is an objective probability distribution over the intentions that it is possible for the postulated necessarily existent agent to form, then it seems clear that naturalists can quite reasonably suppose that the fourth scheme is, on these grounds alone, inferior to each of the first three schemes. On the other hand, if it is to be supposed that there is an objective probability distribution over the intentions that it is possible for the postulated necessarily existent agent to form, then it seems to me that it is no less clear that naturalists can quite reasonably suppose that the fourth scheme does worse in point of satisfying the combined desiderata for theory choice than do any of the first three schemes. Of course, in this second case, I don't say that one could not reasonably prefer the fourth scheme to each of the first three schemes, all things considered; but I do think that one would surely need to appeal to other considerations, beyond those that are invoked in the context of cosmological arguments.

First loose end. O'Connor makes much of the point that his opponents will likely focus on explanations of contrastive facts, i.e., explanations of facts of the form “C rather than C*,” where C* is a possible but rejected alternative to C. For example, he writes as follows:

> Let us suppose that there is no explanation for many contrastive truths. That is, there is an explanation for why C obtained, a reason that ultimately involves the **(p.185)** necessary being's causal activity as guided by particular reasons, but there is no explanation for why C obtained rather than C*, a possible state of affairs for which there are also motivating considerations. On a natural reading of the question ‘Why did C obtain rather than C*?’ what is being asked for is a factor that ‘tipped the scales’ in favour of C, a factor the obtaining of which ruled out the occurrence of C. But once we analyse the question in this way, we see that it amounts to little more than a demand for an account that renders the actual state of affairs inevitable. But that there is no such account to be given, owing to its contingent origins, does not entail that explanation is absent for what has actually obtained: the concrete reality of persisting particulars interacting and undergoing change through time.[28]

But, if what I have said is right, then there is clearly a sense in which it is quite appropriate to focus on explanations of contrastive facts: for, while there is a sense in which one can give a “complete” explanation of contrastive facts by way of appeals to objective probability distributions, there is no sense in which one can give a “complete” explanation of contrastive facts that appeals neither to counterfactual dependence nor to objective probability distributions. What is at issue is not an improper demand for accounts that render the actual state of affairs inevitable; rather, it is a perfectly respectable demand that we be told exactly how counterfactual dependence and objective probability distribution combine to account for the actual state of affairs.

Second loose end. As we noted in the previous section, O'Connor claims that the fourth scheme, unlike the third, is able to satisfy the desideratum that a scheme should provide an explanation for every fact other than those for which there is an explanation of why there can be no explanation of those facts. The problem that he finds for the third scheme is that, even if the universe has no explanation, it could have had one, as there could have been an equally contingent being that caused it to be. But, equally, even if there is no explanation of the intentions that are generated by the necessarily existing agent, because there is no objective probability distribution over those intentions, it seems that there could have been such a distribution. If that's right, then what I take to be O'Connor's preferred interpretation of his fourth scheme does not score better on the desideratum at issue.

Third loose end. I take it that O'Connor's account of agency is just plain false if applied to human agents. The intentions that we generate do depend upon our prior states; what O'Connor calls the “mechanistic” model of agency is, indeed, correct. Of course, O'Connor disagrees: he thinks that the “mechanistic” model of agency is “a mistake” (80).[29]

(p.186) But neither O'Connor nor I make arguments that directly depend upon the presupposed truth or falsity of the “mechanistic” model of agency as it applies to human agents. It would take us very far afield to try to pursue this matter further.

VIII. Concluding Remarks

In my discussion of the existence stage of O'Connor's cosmological argument, I have neither examined nor challenged the views that O'Connor defends in the first part of his book (“The Explanatory Role of Necessity”). While other naturalists might well wish to challenge some of the controversial theses that O'Connor defends in that first part of his book, I have been happy to let all of these theses stand, at least for the sake of argument. My project here has been to explore how O'Connor's cosmological argument fares if we grant him as much as we can of his underlying metaphysical assumptions. My conclusion is that, even given this concessionary approach, the argument does not fare particularly well.

Roughly speaking, there are three clear naturalist alternatives to O'Connor's theistic scheme: infinite regress, brute fact, and naturalistic necessary origin. Naturalists have good reason to think that, if there is infinite regress, then naturalism is preferable to theism. Moreover, naturalists have good reason to think that either brute fact or naturalistic necessary origin is preferable to theism. Consequently, naturalists have good reason to reject theism—here understood to mean “a theistic explanation of the origins of the universe”—on the grounds that, no matter which naturalistic alternative is correct, it trumps the theistic explanation. Naturalists may well think—as I do—that there are no particularly good grounds for choosing between the three naturalistic alternatives: there is nothing that speaks strongly for or against any of these schemes. However, so long as each of the undefeated naturalist schemes trumps theism, naturalists can be justified in their naturalism without making a choice.

References

Bibliography references:

Kripke, S. (1981), *Naming and Necessity* (Oxford: Blackwell).

O'Connor, T. (2008), *Theism and Ultimate Explanation* (Oxford: Blackwell).

Pruss, A. (2006), *The Principle of Sufficient Reason* (Cambridge: Cambridge University Press)

Notes:

(1) O'Connor (2008).

(2) Ibid. (ix).

(3) Ibid. (xii).

(4) Ibid. (110).

(5) Ibid. (143).

(6) O'Connor (2008: xii, 65).

(7) Ibid. (66, 69).

(8) Ibid. (65).

(9) Ibid. (xii).

(10) O'Connor (2008: 65).

(11) Ibid. (72–3).

(12) O'Connor (2008: 67).

(13) Ibid. (87–92).

(14) Ibid. (xii).

(15) O'Connor (2008: 74).

(16) Pruss (2006: 44).

(17) Ibid. (75).

(18) O'Connor (2008: 69–70).

(19) O'Connor (2008: 75).

(20) Ibid. (84).

(21) See Kripke (1981: 110–14).

(22) O'Connor (2008: 82).

(23) Ibid. (80).

(24) Ibid. (82–3).

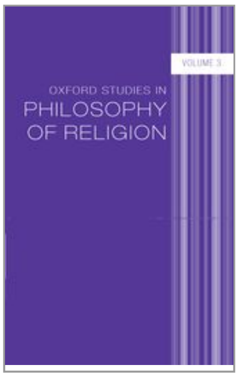
(25) Ibid. (74).

(26) O'Connor (2008: 85).

(27) Ibid. (83).

(28) O'Connor (2008: 84).

(29) Ibid. (80).



Oxford Studies in Philosophy of Religion Volume 3
Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

Evolution without Naturalism

Elliott Sober

DOI:10.1093/acprof:oso/9780199603213.003.0010

Abstract and Keywords

God and numbers provide two challenges to metaphysical naturalism–the former if God exists and is a supernatural being, the latter if numbers exist and mathematical Platonism is true. Evolutionary theory is often described as having a commitment to naturalism, but this is doubly wrong. The theory is neutral on the question of whether God exists and mathematical evolutionary theory entails that numbers exist. The chapter develops the point about theistic neutrality by considering what evolutionary biologists mean when they say that mutations are “unguided.” Evolutionary theory is logically compatible with deism and also with various interventionist theologies. In connection with mathematical evolutionary theory, the chapter discusses and criticizes the indispensability argument for mathematical Platonism developed by Quine and Putnam. The chapter criticizes their epistemological holism by considering ideas in Bayesian confirmation theory.

Keywords: Bayesianism, confirmation, holism, indispensability, naturalism, Platonism, unguided mutation

> “…any confusion between the ideas *suggested* by science and science itself must be carefully avoided.” Jacques Monod, *Chance and Necessity*, xiii.

Does evolutionary theory have implications about the existence of supernatural entities? This question concerns the *logical relationships* that hold between the theory of evolution and different bits of metaphysics. There is a distinct question that I also want to address; it is *epistemological* in character. Does the evidence we have for evolutionary theory also provide evidence concerning the existence of supernatural entities?

An affirmative answer to the logical question would entail an affirmative answer to the epistemological question if the principle in confirmation theory that Carl Hempel called the special consequence condition were true:

> *The special consequence condition*: If an observation report confirms a hypothesis H, then it also confirms every consequence of H.[2]

According to this principle, if evolutionary theory has metaphysical implications, then whatever confirms evolutionary theory also must confirm those metaphysical implications. But the special consequence is false. Here's a simple example that illustrates why. You are playing poker and would dearly like to know whether the card you are about to be dealt will be the Jack of Hearts. The dealer is a bit careless and so you catch a glimpse of the underside of the card on top of the deck before it is dealt to you. You see that it is red. The fact that it is red confirms the hypothesis that the card is the Jack of Hearts, and the hypothesis that it is the Jack of Hearts entails that the card will be a Jack. **(p. 188)** However, the fact that the card is red does not confirm the hypothesis that the card will be a Jack.[3] Bayesians gloss these facts by understanding confirmation in terms of probability raising:

> *The Bayesian theory of confirmation*: O confirms H if and only if $Pr(H|O) \rangle Pr(H)$.

The general reason why Bayesianism is incompatible with the special consequence condition is depicted in Figure 1.[4]

Although you don't have to be a Bayesian to see that the special consequence condition is mistaken, the Bayesian analysis of why it fails in the poker example suggests a larger pattern. The hypothesis of interest is a conjunction; it says that the next card is a Jack *and* that it is a heart. The observation that the card is red raises the probability of the second conjunct and also raises the probability of the whole conjunction; however, it does not raise the probability of the first conjunct. Perhaps evolutionary theory is a conjunction of metaphysical and empirical elements; observations may confirm the theory by confirming its empirical elements but without confirming what is metaphysical. I offer this as a thought to ponder, not as a thesis now to be embraced. Beware! There are reasons to hesitate. Maybe the theory, properly construed, has no metaphysical elements at all. Or maybe

(p.189) the metaphysical elements and the empirical elements can't be separated in this neat way. Let us see.

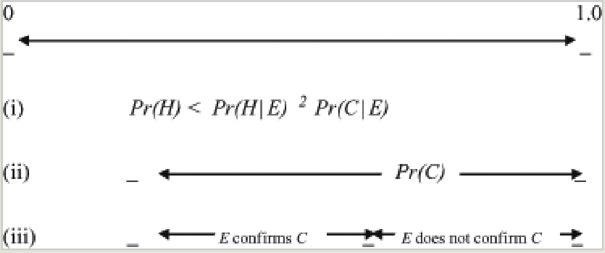


Fig. 1. Why Bayesianism entails that the special consequence condition is false. If *E* confirms *H*, the inequality between the first and second terms in (i) holds. If *H* entails *C*, the inequality holds between the second and third term in (i). Given just these two facts, the value o f *Pr(C)* could be anywhere in the interval shown in (ii). Where *Pr(C)* is located determines whether *E* confirms *C*, as shown in (iii).

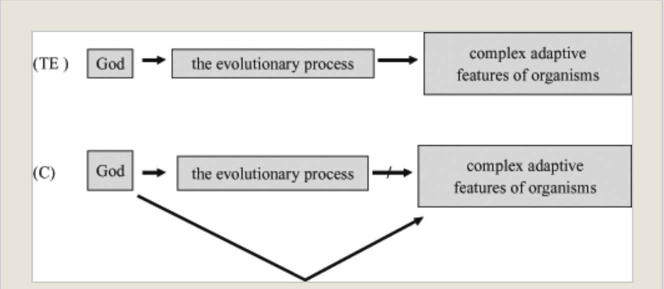
In what follows, my two main metaphysical topics are God and numbers. I'll begin with the familiar general point that the existence of God is compatible with evolutionary theory and then turn to the more specific question of what evolutionists mean when they say that mutations are unguided. Doesn't this entail that there is no God who guides the evolutionary process? I'll then consider whether theistic evolutionists must be deists. Must they hold that God starts the universe in motion and never intervenes in what happens after that? Although I will focus mainly on how theistic evolutionism should be formulated, I hope the points I'll make will also be relevant to how agnostics and atheists should think about scientific theories. The upshot is that evolutionary theory is neutral on one question about naturalism—the question of whether a supernatural deity exists. Numbers, my second main topic, raise a different set of issues, since mathematical evolutionary theory entails that numbers exist. If Platonism is the right philosophy of mathematics, then evolutionary theory (like other mathematized sciences) entails that naturalism is false. Returning to the inadequacy of the special consequence condition, I'll argue that the evidence we have for evolutionary theory isn't evidence for the existence of numbers. Here I'll criticize the Quine/Putnam indispensability argument and defend the anti-holistic thesis that theories typically have many parts and scientists are within their rights to hold that these parts sometimes differ in their epistemic status. My overall goal in this essay is to challenge the influential idea that the opposition between evolutionary theory and creationism is a clash between two philosophies—naturalism and supernaturalism. In one respect, evolutionary theory is uncommitted as to whether naturalism is true; in another respect, it is anti-naturalistic.

I. God

Creationists maintain that the theory of evolution entails that there is no God. If they are right, then the theory has metaphysical implications. Atheistic evolutionists (e.g., Dennett 1995 and Provine 1989) often agree with creationists on this point. The conditional "if evolutionary theory is true, then there is no God" is therefore common ground. Where creationists have their *modus tollens*, these evolutionists have their *modus ponens*.

Both sides are wrong. Theistic evolutionism is a logically consistent position (Ruse 2000; Sober 2008). This is the idea that God uses the evolutionary **(p. 190)** process to make organisms. The difference between creationism and theistic evolutionism is depicted in Figure 2. Creationists don't limit themselves to the assertion that there is a God who is responsible for the living things that we observe. They additionally maintain that the evolutionary process is inherently incapable of producing the complex adaptations that organisms have and that these are due to God's *direct* intervention in nature. According to theistic evolutionism, God produces organisms *indirectly*, by setting the evolutionary process in motion.

In saying that theistic evolutionism is logically consistent, I am not saying that it is plausible or true. I'm merely saying that it isn't contradictory. Evolutionary theory is silent on the question of whether God exists. Even if you think the theory knocks the wind from the sails of the argument from design,[5] you still need to consider the fact that there are other arguments for the existence of God. And even if you think that none of these arguments is rationally persuasive, you still need to consider whether belief in God must be based on evidence. This is the old debate between William James's (1897) "Will to Believe" and W. K. Clifford's (1877) "The Ethics of Belief," which Alvin Plantinga has engaged via his thesis that belief in God is "properly basic."[6] Evolutionary theory has no implications about the cosmological **(p.191)** argument or the ontological argument, nor does it say whether you are entitled to believe in God even if you can offer no good argument that such a being exists. Atheistic evolutionists may scoff at these other arguments and at those who believe in God while admitting that they can offer no compelling argument that there is such a being. But this is not the theory of evolution talking.


(TE)
God
the evolutionary process
complex adaptive features of organisms
(C)
God
the evolutionary process
complex adaptive features of organisms

Although evolutionary theory is silent on the question of whether there is a God, it is not neutral with respect to logically stronger hypotheses about God. Consider, for example, the statement:

> Life appeared on earth about 10,000 years ago due to divine intervention.

Fig. 2. Theistic evolutionism (TE) says that God made organisms and their complex adaptive features by setting the evolutionary process in motion. Creationism (C) says that evolutionary process is incapable of producing the complex adaptive features we observe and that these were the result Gods direct intervention in nature.

This statement *is* inconsistent with evolutionary theory. Does this show that the theory has metaphysical implications? After all, it entails that there is no God of the sort just described. Maybe so, but only in the trivial sense that every statement is thereby shown to have metaphysical implications. Consider the statement "my car is gray." This entails the falsehood of the statement:

> A supernatural intelligent designer ensured that everything in the physical world would fail to be gray.

Is the statement about my car thereby shown to have metaphysical implications? I am inclined to think that it is much more humdrum. In any event, our inquiry is less likely to lapse into triviality if we ask about specific metaphysical propositions and whether evolutionary theory has anything to say about them. The question is not whether evolutionary theory has any implications about metaphysical matters, but whether it has implications, for example, about whether there is a God. And it does not.[7]

II. Undirected Mutation

What is this thing called "evolutionary theory," which theistic evolutionism is able to encompass consistently? It includes the origin of life from nonliving materials by physical processes, the branching genealogical process whose upshot is that current organisms are connected to each other by relations of common ancestry, the random origin of new mutations, and the processes that govern trait evolution within lineages, such as selection and **(p.192)** drift. Among these several propositions, the idea that mutations are "random" may seem to be a sticking point. Random mutation suggests "blind chance," but what God causes does not occur by blind chance. Theistic evolutionism is often thought of as evolution under the guidance of God. How can this be true if mutations are unguided?

When evolutionists say that mutations occur “at random,” they do not mean that mutations are uncaused. Rather, their point is that mutations do not occur *because* they would be useful to the organisms in which they occur. It is well known that there are physical events, like radiation, that influence the probability of mutations. Some of these causes raise mutation rates at many loci. Others are more specific, raising the probability of some mutations but not of others. And there are parts of the genome, so-called “hot spots,” that are especially vulnerable to mutation. Suppose that an organism in a given environment would benefit from having mutation M occur. There are physical facts about the organism and its environment that make this mutation have whatever probability it has. But this beneficial mutation is not more apt to occur because it would be beneficial. The causation goes through the physical facts about the organism and its environment, period. As a first pass, let's consider expressing the idea that mutations are undirected by saying that the physical causes of mutation (which encompass both the physical make-up of the organism and physical facts about its environment) *screen off* the mutation's occurring from whether or not it is beneficial:

$$\begin{aligned}\text{(SO)}\quad & \mathbf{Pr(mutation}\mathit{M}\mathbf{occurs\ in}\mathit{O}\mathbf{at}t_2|\mathbf{E\ at}t_1) = \\ & \mathbf{Pr(mutation}\mathit{M}\mathbf{occurs\ in}\mathit{O}\mathbf{at}t_2|\mathbf{E\ at}t_1\mathbf{and}\mathit{M}\mathbf{would\ be\ beneficial\ to}\mathit{O}\mathbf{at}t_2).\end{aligned}$$

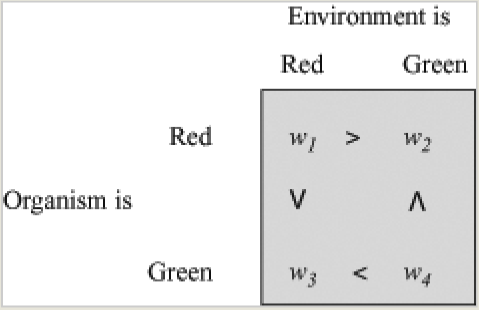
Here E is a complete description of the physical facts about the organism and its environment at the earlier time t_1.

The trouble with SO is that it doesn't capture the thesis that there is no *physical mechanism* (either inside organisms or outside of them) that detects which mutations would be beneficial and causes those mutations to occur. The SO equality merely expresses the thesis that the physical facts at t_1 are *causally complete* with respect to the occurrence of mutations at t_2.[8] And why isn't it a “physical fact” that a mutation would help an organism to survive and reproduce? Some hints about how this formulation can be improved **(p.193)** may be found in Lenski and Mittler's description of what it means for a mutation to be directed:

> We define as directed a mutation that occurs at a higher rate specifically when (and even because) it is advantageous to the organism, whereas comparable increases in rate do not occur either (i) in the same environment for similar mutations that are not advantageous or (ii) for the same mutation in similar environments where it is not advantageous.[9]

The basic idea here can be understood in terms of a simple example. Consider a blue organism that might live in a green or a red environment and suppose that protective coloration would provide it with an adaptive advantage; being green in a green environment and being red in a red environment would be good, but being green in a red environment and being red in a green environment would be bad. This means that there are two fitness peaks and two valleys in the table shown in Figure 3.

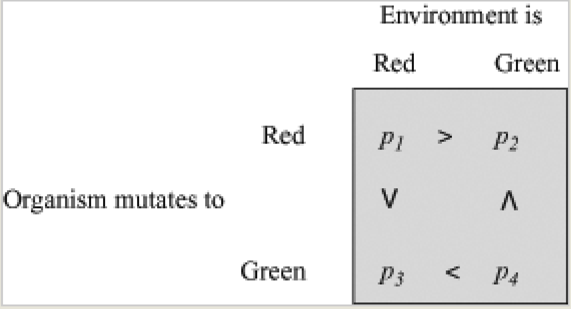
Let's consider two possible mutations that these blue organisms might experience; one will cause its offspring to be green and the other will cause its offspring to be red. These two mutations are directed only if inequalities among mutation probabilities *mirror* inequalities among the fitnesses given in Figure 3. The pattern is shown in Figure 4; peaks and valleys in fitnesses match peaks and valleys in mutation probabilities. I've introduced two small changes in Lenski and Mittler's formulation. First, I use the comparative concept of one mutation's *being more beneficial* than another rather than the absolute concept of a mutation's *being beneficial*. I've also used probabilistic inequalities rather than rate inequalities; this is **(p.194)** because I think of rates as observations that provide evidence about probabilities, and it is the latter that characterize the relevant mechanisms. I trust it is clear how the pattern of mutation probabilities mirroring fitnesses can be generalized to more than two environments and more than two mutations.[10] Notice that the concept defined in Figures 3 and 4 is specific to one species of organism, a pair of mutations, and a pair of environments. This leaves it open that mirroring might hold in some species, but not in others. It also allows that there might be mirroring for some of the mutations a given species experiences but not for others; it also is possible that mirroring might hold in some environments, but not others. Although the screening-off thesis SO is totally general, questions about the mirroring relation must inevitably be addressed piecemeal. And whereas SO is settled by a general commitment to a physicalistic metaphysics, questions about mirroring are not settled by philosophy, but need to be addressed via specific empirical (preferably experimental) results.



		Environment is		
		Red		Green
	Red	w_1	$>$	w_2
Organism is		$\vee$		$\wedge$
	Green	w_3	$<$	w_4

Fig. 3. Four fitness values and their relationships. If protective coloration is advantageous to an organism, there are two fitness peaks and two valleys.

It would not be very surprising if an organism, a pair of environments, and a pair of mutations were discovered that exhibit this mirroring relationship. Surely we can cherry-pick from a large data set to find a few examples of this sort. Biologists would not only decline to conclude from this that **(p.195)** mutations are *generally* directed; I think they'd also decline to conclude that *these* mirroring mutations are directed. Mirroring is a necessary condition for a set of mutations to be directed; it isn't sufficient. What is missing from the relationship of mirroring is the word "because." What needs to be added is that more beneficial mutations have higher probabilities than ones that are less beneficial *because* this asymmetry is beneficial.[11] It is here that the concept of natural selection makes its appearance. *Prima facie*, it seems clear that a mechanism that increases the frequency of advantageous mutations and reduces the frequency of deleterious mutations would be favored by natural selection.[12] Unfortunately, counting mutation events in laboratory populations of micro-organisms like *E. coli* is difficult. What one observes, in the first instance, is the *traits* of organisms; whether an organism's trait is due to a new mutation or is part of the standing variation in the population is often difficult to discern. This discrimination is critical if one wants to determine whether additional mutations occurred *after* a new selective environment was imposed. The history of this subject is one in which biologists have carefully devised experiments, but other biologists then uncover subtle flaws in the experiment design or think of new hypotheses that account for the results that do not require the hypothesis of directed mutation. It is to be expected that new and better experiments will be made. But, for now, the hypothesis that mutations are guided should be viewed with skepticism. True, the ability to direct mutations would be favored by natural selection if it were present in the population. But that does not mean that it ever was present as a variant. Selection can't cause a trait to increase in frequency unless it is exemplified in the first place.[13]



		Environment is		
		Red		Green
	Red	p_1	>	p_2
Organism mutates to		∨		∧
	Green	p_3	<	p_4

Fig. 4. Four mutation probabilities. If the mutaions are directed, inequalities among probabilities must mirror the inequalities among the fitness relationships shown in Figure 3.

The thesis that mutations are undirected is sometimes presented as a rock-bottom "philosophical" commitment of evolutionary biology, akin to **(p.196)** materialism and just as central to the scientific worldview. This is a mistake. The thesis is certainly central to biology, but it is important to see that it has nothing to do with materialism or theism. It is like the thesis that an organism can't synthesize vitamin *D* from sunlight or can't regenerate lost limbs. It is an empirical question whether the mirroring relation obtains for a given organism with respect to a given set of mutations in a given range of environments. If it does, that is something evolutionary biology needs to explain. Mutation rates vary among and within species and it is a good question why they do so. Explaining patterns of variation in mutation rates is a problem *within* evolutionary biology, not a *challenge to* evolutionary biology.

I conclude that the fact that mutations are undirected is not a problem for theistic evolutionism. Maybe God arranged for mutations to be undirected. And if some mutations do turn out to be directed in the sense described here, that is no threat to atheism. Atheism has no more of a stake in mutations being undirected than it has in organisms being unable to synthesize vitamin *D* from sunlight.

Ernst Mayr (1961) emphasized the importance of distinguishing proximate from ultimate causation. Mayr's point can be illustrated by considering the question of why sunflowers turn towards the sun. One answer to this question cites the proximate mechanism inside each sunflower that causes it to do so; the other points to the reason the trait evolved, presumably natural selection. There is no conflict between these two "levels" of explanation. The theistic evolutionist should embrace Mayr's pair of concepts and add a third, maintaining that there is, in addition, a still more distal level of divine causation. God can direct the evolutionary process in an ultimate sense, though mutations are undirected in a proximate sense. Biology says nothing about the former and theism says nothing about the latter.

III. Evolution and Hidden Variables

If theistic evolutionism is correct in its claim that the only pathway from God to the adaptations we observe passes through the evolutionary process, this means (if there is no action at a distance) that the evolutionary process *screens off* God from evolutionary outcomes:

$$\Pr(\text{adaptations we observe} \mid \text{evolutionary process}) =$$

$$\Pr(\text{adaptations we observe} \mid \text{evolutionary process and God's choices}).$$

This should not be taken to mean that evolutionary *theory* does the screening off. The theory is probabilistic; it does not say of itself that it is *causally complete*. The theory is consistent with there being *hidden variables, natural* ***(p.197)*** *or* supernatural. For example, the theory is consistent with determinism at the micro-physical level.[14] This means that even if a probabilistic evolutionary model says that some possible outcome has a non-extreme probability, a complete physical model may say that the outcome's probability is 0 or 1. Evolutionary theory is silent on this question about physics.

To get a feeling for the distinction between (a complete description of) the evolutionary process's screening off and evolutionary theory's doing so, consider the process of Mendelian inheritance. It is generally thought that this process is Markovian—that grandparents influence the genotypes of their grandchildren only by way of influencing their children. There is no action at a temporal distance in genetic transmission. However, this fact about the process does not mean that there can't be models in population genetics in which the state of generation 3 depends not just on the state of generation 2 but on that of generation 1 as well. For example, in the theory of inbreeding, when all matings are between siblings, the heterozygosity (H_t) in generation t depends on the heterozygosity in the previous two generations, $t-1$ and $t-2$:[15]

$$H_t = (1/2)(H_{t-1}) + (1/4)(H_{t-2})$$

The *process* is Markovian, but *models* of the process can fail to be; this just means that the models are causally incomplete.

It is consistent for theistic evolutionists to claim that there is more going on in the evolutionary process than is dreamt of in evolutionary biology, just as a friend of determinism can consistently maintain that there is more involved in the process of flipping a coin than is described in a simple probabilistic model:

$$\mathrm{Pr}(\text{coin lands heads at} t_2 | \text{the coin is tossed at} t_1) = 0.5.$$

This probability statement is entirely consistent with the truth of another:

$$\mathrm{Pr}(\text{the coin lands heads at} t_2 | \text{a complete description of the initial conditions at} t_1) = 0 \text{ or } 1.$$

That both can be true is no more puzzling than the fact that *X* doesn't entail *Z* while *X&Y* does. Asking what the true probability is of *Z* makes no sense. The question is incomplete; the proper question is—what is the probability of *Z*, given this or that other proposition?

(p.198) A probabilistic model of coin tossing is consistent with the thesis that the system is deterministic. If determinism is true, there are hidden variables, not represented in the probability model, which turn all the probabilities into 0s and 1s when their values are taken into account. In just the same way, a probabilistic model of the evolutionary process is consistent with the thesis that the process is deterministic. If determinism is true, there are hidden variables that affect the evolutionary process. Evolutionary theory says nothing about whether such hidden variables exist. It therefore says nothing about whether there are supernatural hidden variables.

Does this mean that evolutionists overstep what they know when they say that evolution is an undirected process? No more than experts on gambling devices overreach themselves when they say that a roulette wheel obeys a certain probability model. In both cases, there is evidence that favors the probability model in question over competitors—*this* is what is behind the scientific pronouncement. Those who hold that everything that happens is due to the sustaining power of God can accept what experts say about gambling devices; they should take the same view of what evolutionists say about the evolutionary process.

Theistic evolutionists can of course be deists, holding that God starts the universe in motion and then forever after declines to intervene. But there is no contradiction in their embracing a more active God whose post-Creation interventions fly under the radar of evolutionary biology. Divine intervention isn't part of science, but the theory of evolution does not entail that none occurs.[16]

In this light, it is worth considering a passage from one of Charles Darwin's letters to Charles Lyell:

> Of course, it may be said, when you kick a stone, or a leaf falls from a tree, that it was ordained before the foundations of the world were laid, exactly where that stone or leaf should lie. In this sense the subject has no interest for me.[17]

Whether or not Darwin meant this last comment in this way, I interpret it to mean: *my theory has nothing to say about this matter.* The theory does not rule this out, however much Darwin the man may have been skeptical of this proposition.

(p.199) IV. Ockham's Razor and Theistic Evolutionism

Creationists have been hostile to theistic evolutionism. The same is true of those who prefer to call themselves Intelligent Design Theorists. Here is what William Dembski says about the position:

> …intelligent design is incompatible with what typically is meant by theistic evolution. Theistic evolution takes the Darwinian picture of the biological world and baptizes it, identifying this picture with the way God created life. When boiled down to its scientific content, however, theistic evolutionism is no different from atheistic evolutionism, treating only undirected natural processes in the origin and development of life…
>
> If theistic evolution finds no solace from intelligent design, neither does it find solace from the Darwinian establishment. For the Darwinian establishment the theism in theistic evolutionism is superfluous. For the hard-core naturalist, theistic evolution at best includes God as an unnecessary rider in an otherwise purely naturalistic account of life. Thus by Ockham's razor, since God is an unnecessary rider in our understanding of the natural world, theistic evolution ought to dispense with all talk of God outright and get rid of the useless adjective *theistic*. This is the received view within the Darwinian establishment.[18]

This remark does not provide an *argument* that theistic evolutionism is false, but expresses a *motive* that Dembski and others have for wanting creationism to be the only game in town.

Dembski and other creationists are worried that the hypothesis that God exists will be rendered superfluous by science. If the God hypothesis isn't needed to explain the adaptive features of organisms, there is no biological reason to think that there is such a being. If the same thing happens in the other sciences, there will be no scientific reason to think that there is a God. By “reason,” I mean evidence. Ockham's razor then seems to tell us to slice away the proposition “God exists” from our corpus of beliefs. This slicing away will result in agnosticism or atheism, depending on how the principle of parsimony is formulated.

There are two ways in which a theistic evolutionist might answer this argument. The first is to point out that all scientific explanations rely on a description of a system's initial conditions.[19] This is true in physics just as much as it is true in biology. If the universe isn't infinitely old, there is such a thing as the initial state of the universe, and no scientific explanation clarifies the character of that initial state. Does the God hypothesis thereby explain something that would otherwise be inexplicable? Atheists may **(p.200)** complain that this initial state does not need to be explained, or that the God hypothesis in fact offers no genuine explanation, or that the hypothesis that God caused the initial state is untestable. All these options are worth considering. Evolutionary theory has nothing to say about them.

Theistic evolutionists can reply to the argument from Ockham's razor in a second way. They can point out that the principle of parsimony is part of the scientific method, but that they don't think that one needs to justify belief in God by appeal to the scientific method. This is the simple idea that belief in God is a matter of faith, not evidence. Atheists will probably regard this reply as an abdication, perhaps taking the side of W. D. Clifford in his disagreement with Williams James. The debate between these two philosophies is—no surprise—*philosophical*, not *biological*. The theory of evolution has nothing to say about this, either.

V. Numbers

The debate over evolutionary theory versus creationism is often assimilated to the conflict between two philosophies—naturalism and supernaturalism. Both are metaphysical doctrines. The former holds that the only objects that exist are objects *in* nature; the latter denies this, claiming that at least some objects exist *outside* of nature—they are *super*natural. Nature, for our purposes, can be understood as the totality of things located in space and time. We've just seen that when God is the supernatural entity in question, it is a mistake to think that evolutionary theory is committed to naturalism in the metaphysical sense just described. But there is a second issue about metaphysical naturalism; it comes from the philosophy of mathematics, not the philosophy of religion, and it raises different issues. Here a case can be made, not that evolutionary theory is *silent* on the issue of naturalism, but that it actually entails that metaphysical naturalism is *false*.

Mathematical Platonism says that mathematical statements describe objects that exist outside of space and time. These objects are said to exist independently of human thought and language. Even though there would be no *names* for numbers *or mental representations* of mathematical facts if there were no intelligent beings, the numbers themselves would still exist and mathematical propositions would still be true. Not only are mathematical truths independent of whether there are beings with mind and language; they also are independent of whether there are any material objects. For example, the prime number theorem—that there exist an infinite number of prime numbers—does not depend for its truth on how much mass there is in the universe or on how that mass is distributed. In view of the fact that **(p.201)** mathematical propositions are true independent of what is going on *inside of* nature, it is natural to think that those propositions describe objects that exist *outside of* space and time. I don't want to give the impression that mathematical Platonism is *obviously* true. Some serious challenges have been raised. For example, Paul Benacceraf has argued that if mathematical propositions are true or false in virtue of facts about such Platonic entities, that it would be impossible for human beings to have mathematical knowledge.[20] But we clearly do have some mathematical knowledge. And Hartry Field has argued that some portions of mathematics can be reinterpreted in such a way that their subject matter consists entirely of things that exist inside of space and time.[21] He also has suggested that the parts of mathematics that are irreducibly Platonistic in their commitments should be regarded as useful fictions. These challenges are serious and interesting, but they should not lead us to conclude that mathematical Platonism is obviously mistaken. If this metaphysics is false, it is not easy to refute, and the alternative positions have problems of their own.[22]

What has this question in the philosophy of mathematics to do with evolutionary theory and its metaphysical commitments? If mathematical Platonism is true as an account of *pure* mathematics, then it also is true as an account of *applied* mathematics.

Evolutionary theory became a mathematical subject in the twentieth century, thanks initially to work in population genetics. Mathematical evolutionary theory makes claims about genes, organisms, and populations, but it also makes claims about numbers. In this respect it is just like mathematical physics. Consider, for example, a simple one-locus two-allele model of natural selection in which heterozygote superiority entails that the population will evolve to an equilibrium gene frequency in which both alleles are maintained. The mathematical content and Platonistic commitments of this model become clearer when we formalize this implication of the model as follows:

> For any diploid population, and for any locus in that population in which there are two alleles (*A* and *a*), and for any frequency f ($0 \langle f \langle 1$) that the *A* allele has in the population at time $t = 0$, and for any three fitness values p_1, p_2, and p_3 that the three genotypes (*AA*, *Aa*, and *aa*) might have, if $w(AA)=p_1 \langle w(Aa)=p_2 \rangle w(aa)=p_3$, then there exists an equilibrium frequency e (where $0 \langle e \langle 1$) such that $f \to e$ as $t \to \infty$.

(p.202) Here the fitness value of a genotype is the probability that an organism with that genotype has of surviving from egg to adult. This model quantifies over numbers—probabilities and frequencies. If Platonism is the right philosophy of mathematics, evolutionary theory is incompatible with naturalism.

Although mathematized evolutionary theory entails that there are numbers, the empirical evidence for evolutionary hypotheses does not provide evidence that numbers exist. Here we are back to the earlier point about the falsity of Hempel's special consequence condition. One way to see what is going on here is to think about evidence in terms of the Law of Likelihood:[23]

> *The Law of Likelihood*. Evidence E favors hypothesis H_1 over hypothesis H_2 precisely when $Pr(E|H_1) \rangle Pr(E|H_2)$.[24]

The Law of Likelihood is central to Bayesianism; for the Bayesian, it describes the one and only avenue through which observational evidence can modify one's degrees of belief. The odds version of Bayes's Theorem (Figure 5) shows why. Notice that the ratio of the posterior probabilities can differ from the ratio of priors only if the likelihoods of the hypotheses are different. And the more the likelihood ratio deviates from unity, the more the ratio of posterior probabilities will differ from the ratio of priors. The likelihood ratio represents the power that the observations have to transform one's degrees of belief.[25]

$$\underbrace{\frac{\mathrm{Pr}(H_1\_E)}{\mathrm{Pr}(H_2\_E)}}_{\text{ratio of posterior probabilities}} = \underbrace{\frac{Pr(E\_H_1)}{Pr(E\_H_2)}}_{\text{likelihood ratio}} \times \underbrace{\frac{Pr(H_1)}{Pr(H_2)}}_{\text{ratio of prior probabilities}}.$$

Fig. 5. The odds version of Bayes Theorem.

(p.203) The Law of Likelihood suggests the general thesis that testing in science is typically *contrastive.*[26] To test a hypothesis *H*, you must test *H against* some alternative. There are two possible exceptions to this pattern; each arises when a hypothesis and a set of observations are deductively related. If *H* entails *E* and you see that *E* is false, it follows that *H* is also false, there being no need to consider alternatives to *H* along the way. And if *E* entails that *H* is true, and you see that *E* is true, it follows that *H* is true: no need for contrasts here, either. Neither of these patterns has much purchase in the world of scientific testing. The uselessness of the second pattern is especially clear—rarely are the hypotheses that scientists want to test deducible from the observations. It may be thought that a point due to Pierre Duhem and Willard van Orman Quine vitiates the interest of the first pattern.[27] Their logical point is that, typically, you need auxiliary propositions to bring the hypothesis you wish to test in contact with observations. It isn't *H* that entails *E*, but the conjunction *H&A* that does so. The point is right, but it doesn't show that testing must be contrastive. If you know that *A* is true, and if the conjunction *H&A* entails *E*, and *E* turns out to be false, you can deduce that *H* is false; hypotheses alternative to *H* play no role. The real reason that testing almost always has to be contrastive is that hypotheses, with independently supported auxiliary assumptions in place, almost always confer non-extreme probabilities on the evidence. *H* doesn't entail *E*, but neither does *H&A.*

Not only is testing almost always contrastive. In addition, testing in science typically pits different empirical theories against each other that make use of the same mathematical resources. This shared mathematical background isn't tested, but assumed. When evidence discriminates between two empirical theories H_1 and H_2 that use the same mathematical framework *M*, the usual pattern is that the evidence says nothing about whether *M* is true. This is because *M* has not been tested against an alternative mathematics *M**; in fact, the usual situation is that no such alternative mathematics has even been formulated.

Although the Law of Likelihood represents testing as a contrastive enterprise, this point about testing is not unique to “likelihoodism” or Bayesianism. Non-Bayesian inferential frameworks are used in science and they have the same character. For example, model selection theory is a part of statistics in which the hypotheses under evaluation contain adjustable parameters (this is what the word “model” means in this branch of statistics). **(p.204)** Consider two evolutionary hypotheses concerning the two alleles *A* and *B* that exist at a genetic locus in a population that you are studying. One hypothesis says that the genes of interest differ in fitness and that their fitnesses are independent of their frequencies:

$$\text{(FIS)Thereexitnumbers}w_1\text{and}w_2\text{suchthatw(A)} = w_1\text{andw(B)} = w_2$$

The other model says that each of the alleles has a frequency dependent fitness, where an allele's fitness value linearly depends on how common or rare it is:

$$\text{(FDS)There exist numbers}m_1, m_2, b_1, \text{and b2 such that for any frequency}f\text{of the}A\text{allele, w}(A) = m_1 f + b$$

Instances of the two models (with their adjustable parameters replaced by constants) are depicted in Figure 6.

One influential idea in model selection theory is called the Akaike Information Criterion (AIC). It scores a model by using data to obtain maximum likelihood estimates of its adjustable parameters. Models like FIS and FDS, by virtue of their existential quantifiers, are infinite disjunctions; each disjunct is a “member” of the model. By fitting a model M to the data, one obtains the proposition $L(M)$, which is the member of the model that makes the data most probable. AIC takes the log-likelihood of $L(M)$ and then subtracts a penalty for complexity (= the number of adjustable parameters):

$$\text{The AIC score of model} M = \log\{\Pr[\text{data}|\text{L(M)}]\} - \text{k}.$$

Notice that AIC does not compute the likelihood of the model M, but of the fitted model $L(M)$. And it doesn't use the likelihood of $L(M)$ as the model's score, but instead subtracts k; models that differ in their complexity **(p.205)** have different penalties imposed. What is of interest about AIC is not a model's *absolute* score, but *comparisons* of the scores of different models.[28] AIC and the Law of Likelihood are different approaches to representing what the evidence is saying,[29] but both are contrastive. In our example, the models FIS and FDS have different numbers of adjustable parameters. Data drawn from the population concerning the birth and death rates found in a sample of organisms allow us to decide which model has the better AIC score. However, the observations that allow us to assign AIC scores to FIS and FDS do not permit us to assign AIC scores to the propositions “numbers exist” and “numbers do not exist.” These propositions are not “models” in the sense of that term used in statistics. According to this body of statistical theory, the data say nothing about whether numbers exist. AIC is a procedure for scoring *models*, but not every consequence of a model is a model.[30]

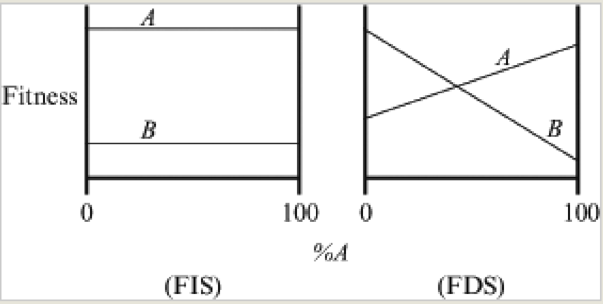


Fig. 6. Instantiations of the frequency independent selection (FIS) and the frequency dependent selection (FDS) models.

I am not here endorsing the false thesis that pure mathematics is never empirically tested. When Joseph Plateau dipped his wires into soapsuds, he was performing an empirical test of various mathematical propositions.[31] The mathematical questions he addressed have the form—what is the surface of least area that is bounded by a given closed curve? See the example depicted in Figure 7. The physics of how soap bubbles form in wire frames allowed Plateau to address his questions experimentally. The same thing happens when mathematicians use computer-assisted proofs[32] and computer simulations;[33] the output on the computer screen provides observational evidence as to whether a given mathematical conjecture is true. Perhaps the use of paper and pencil also should be regarded as an empirical source of information in mathematical reasoning, as should our reliance on the testimony of mathematical experts. Propositions in pure mathematics are sometimes empirically tested. However, the point remains that testing empirical **(p.206)** hypotheses like FIS and FDS that have mathematical commitments typically don't test those mathematical commitments.

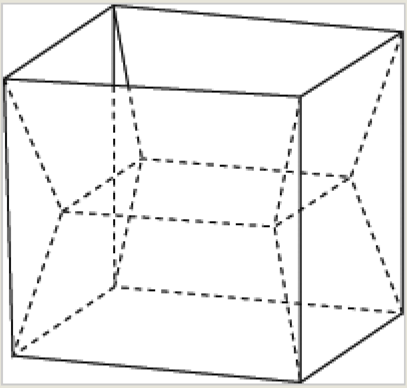


Fig. 7. Onf of Plateau's soap -bubble experiments provided evidence pertaining to the mathematical question of what the surface of least area is that can be inscribed inside a cube.

In the Jack of Hearts example that I gave at the start of this paper, the hypothesis of interest is a conjunction that gets confirmed because one of its conjuncts is confirmed but the other is not. Is this a good model for separating the metaphysical and empirical elements in mathematical evolutionary theory, and, by implication, in other scientific theories? Here it is worth considering M. Friedman's contention that Newtonian physics should not be viewed as a conjunction of pure mathematics and empirical elements:

> …there are fundamental asymmetries in the way in which different elements of this Newtonian synthesis actually function. Consider, for example, the relationship between the mathematics of the calculus and the Newtonian formulation of the laws of mechanics. Newton's second law of motion (in only slightly anachronistic form) says that force equals mass times acceleration, where acceleration is the instantaneous rate of change of velocity (itself the instantaneous rate of change of position); so without the mathematics of the calculus…this second law of motion could not even be formulated or written down, let alone function to describe empirical phenomena. The combination of the calculus plus the laws of motion is not happily viewed, therefore, as a conjunction of elements symmetrically contributing to a single total result. For one element of a conjunction can always be dropped while the second remains with its meaning and truth-value intact. In this case, however, the mathematics of the calculus does not function simply as one more element in a larger conjunction, but rather as a necessary presupposition without which the rest of the putative conjunction has no meaning or truth-value at all. The mathematical **(p. 207)** part of Newton's theory therefore supplies elements of the language or conceptual framework, we might say, within which the rest of the theory is then formulated.[34]

Notice that the conjunction that Friedman considers has all the mathematics in one conjunct. I agree that this isn't plausible, since expressing the empirical content of the theory requires mathematics, just as Friedman says. What I have in mind, rather, is another sort of division, in which ontological commitments to numbers and physical objects are segregated from the rest of the theory:

> (*) Numbers and physical objects exist, and if numbers and physical objects exist, then…

There will be mathematics in the second conjunct and talk of physical objects as well.If we think of the population genetic models FIS and FDS in this format, we can see that testing them against each other really involves testing the second conjunct of FIS against the second conjunct of FDS. I am not suggesting that the first conjunct of (*) exhausts the metaphysics of evolutionary theory; nor am I claiming that “metaphysical elements” always take the form of existence claims.[35] What I hope to have described is a heuristic for isolating some of a theory's metaphysical commitments. The onion can be peeled layer by layer.

VI. Some Objections

I have argued for the following: when an empirical theory makes indispensable use of a body of mathematics, it is a mistake to conclude that the empirical evidence for the theory must also provide empirical evidence for the math. This is a criticism of the indispensability argument for mathematical Platonism made famous by Willard van Orman Quine and Hilary Putnam.[36] My simple point is that the indispensability argument is wrong because the special consequence condition is wrong.[37] The indispensability argument has been criticized on others grounds—that it doesn't explain the obviousness of some mathematical propositions[38] and that it doesn't allow **(p.208)** there to be reasoning internal to pure mathematics that helps one decide which mathematical propositions are true.[39] Although these objections are telling if the indispensability argument purports to identify the *only* pathway by which we gain evidence for mathematical propositions, the argument need not be thought of in that way. These objections do not count against the thesis that indispensability in a well-confirmed empirical theory is *one* such pathway.[40] It is this more modest thesis that I have challenged.

I now want to consider a few criticisms that have been made of my criticism of the indispensability argument.

Hellman concedes that the special consequence condition is often wrong,[41] but maintains that the principle is right for certain key mathematical axioms (e.g., a weak form of the axiom of infinity that holds that an infinite totality exists or is logically possible). They have Bayesian priors that are less than unity, he says, and their probabilities are increased by the fact that they are used in well-confirmed empirical theories. We can return to Figure 1 to see what Hellman's thesis of probability boosting requires. If O confirms H and H entails C, for O to confirm C, it isn't enough that $Pr(C) \langle 1$. Of course, if $Pr(C) \langle Pr(H|\mathrm{O})$, that would suffice for O to confirm C (Kotzen discusses this sufficient condition[42]). The problem is that it isn't clear why agents should arrange their probabilities in this fashion.

Hellman presents a second consideration,[43] which has nothing to do with the Bayesian concept of confirmation as probability raising. Rather, the fact of interest is that probability obeys a rule of weak monotonicity: if H entails C, then $Pr(H \mid X) \leq Pr(C|X)$, no matter what X is. Hellman suggests that weak monotonicity vindicates the indispensability argument. I am skeptical. Notice first that weak monotonicity holds for *any* consequence that H has, whether or not it is mathematical and whether or not it is indispensable; H is related to *each* of its consequences in the way described. Furthermore, although weak monotonicity says that $Pr(H \mid X)$ is a lower bound for $Pr(C|X)$, it also says that $Pr(C|X)$ is an upper bound for $Pr(C|X)$. Hellman seems to be thinking of Bayesian agents who are very confident in the truth of a mathematized theory and extract from this a confidence in the embedded mathematics that is at least as great. But we also should consider Bayesian agents who are not at all confident that numbers exist and who therefore are at least as unconvinced of the mathematized theories that entail that there are such **(p.209)** things. Montonicity says nothing about which thinkers have their heads screwed on better. Many of the scientists I know get impatient when they are asked how sure they are that numbers exist. They regard this as a philosophical question on which they needn't take a stand. When it is pointed out to them that the theories they embrace entail that numbers exist, they reply with surprise that they never thought about that part of the theory. Their confidence, such as it is, is focused on their theories' other parts.

Hellman suggests a third reconstruction of the indispensability argument.[44] Suppose a theory T successfully predicts some observations about which its competitors are silent. Suppose further that T makes use of a strong body of mathematics M, but that its competitors use only the weaker mathematics M'. Hellman thinks that in this case the observations not only confirm T; they also confirm M. The epistemological picture that Hellman is painting here is not in keeping with the law of likelihood, as he notes. In fact, it isn't in keeping with Bayesianism, either. For Bayesians, O confirms H only if $Pr(O|H) \rangle Pr(O|\text{not}H)$; the idea that O confirms H, even though no competitor says anything at all about O, is *impossible*. Although Hellman's first two arguments are couched in a Bayesian framework, this third argument ventures beyond it. More needs to be said about this new epistemology before this proposal for recasting the indispensability argument can be evaluated.

Other critics of my criticism of the indispensability argument also have resisted seeing it through the lens of the Law of Likelihood. Mark Colyvan says that the epistemological framework provided by that principle is a “straightjacket” because it precludes the comparison of empirically equivalent theories on the basis of nonempirical criteria like simplicity and elegance.[45] He is right that the epistemology formulated in an earlier paper[46] has this consequence. The point I would make here is that invoking *nonempirical* criteria to help one evaluate empirically equivalent theories does not rescue the indispensability argument for mathematical Platonism. That argument maintains that the *empirical* evidence for a scientific theory also provides *empirical* evidence for the pure mathematics it uses. If there are *nonempirical* reasons for thinking that numbers exist, that is beside the point. I also see a difference between the Quine/Putnam indispensability argument and what Michael Resnik calls “the pragmatic indispensability argument.”[47] Resnik's argument has nothing to do with a specific theory's empirical confirmation **(p.210)** providing empirical confirmation of the mathematics it embeds. Rather, Resnik takes the fact that we need to use mathematics to do science to show that we are justified in thinking that various mathematical propositions are true. I have no quarrel with Resnik's thesis about pragmatic justification if the point is that scientists aren't being irrational when they assume a mathematical framework. But our being pragmatically justified does not mean that the empirical evidence we have for a theory also provides empirical evidence for the mathematics it embeds.

Philosophers sometimes think of the following argument when they consider the Quine/Putnam indispensability argument:

> One of the theories T_1, T_2,…, T_n is true.
>
> Each theory T_i entails M.
>
> ______________________________
>
> M is true.

If M is common ground among all the competing theories that scientists at a time take seriously, isn't that a good reason for them to think that M is true? After all, you can *deduce* that M is true if you are prepared to assert that the disjunction of these competing theories is true. This is often a big *if*. Scientists do their tests by considering a set of competing hypotheses, but those competitors often fail to be exhaustive. So the first premise of this argument will often be false. In addition, the valid argument displayed above makes no mention of empirical evidence. The fact that the special consequence condition is false applies here, as it did before, and the fact that probability is weakly monotonic is of no help, either.

Colyvan, Hellman, and Resnik[48] each reject a symmetry principle that I recommended in my 1993 article. If a body of pure mathematics gets confirmed by its being used in empirically successful theories, why doesn't it get disconfirmed by its being used in scientific theories that are empirical failures? Think of all the empirically unsuccessful theories that embed the calculus! These three friends of indispensability reject this symmetry principle because they think that the mathematics used in science is confirmable but not disconfirmable. It is worth remembering that, according to Bayesianism, this is *impossible*; *O* confirms *H* if and only if *notO* would disconfirm *H*. Colyvan, Hellman, and Resnik endorse something like the mirror image of Popper's asymmetry thesis that scientific propositions are falsifiable but not verifiable. Even though Bayesianism should not be taken **(p.211)** on faith, the question remains of how a plausible confirmation theory can avoid a symmetry of confirmation and disconfirmation.

Colyvan, Hellman, and Resnik answer by invoking Quine. If a mathematized empirical theory makes successful predictions, it is confirmed and so are its mathematical consequences. If the theory makes unsuccessful predictions, we know that something in the theory needs to be revised. Quine's principle of minimum mutilation then counsels that we revise the nonmathematical parts; we don't reject the math because the math is centrally located in our web of belief. This is not the place to develop a sustained evaluation of these ideas, but I will make a few comments. The first part of this picture needs to be spelled out so that it avoids a wholesale embracing of the special consequence condition. The second uses a dichotomous notion of acceptance, not a notion of evidence that is a matter of degree. And both see testing in terms of the black and white of predictive success and failure. Many have found this Quinean picture suggestive; it remains to be seen whether a plausible and precise epistemology can be developed along these lines.

Colyvan defends the idea that evidence for a theory is evidence for all the theory's parts by way of an intriguing analogy.[49] He suggests that a scientific theory is like a computer program. When a computer program "works" in a series of tests, this is evidence that all its parts are bug-free. Similarly, when a scientific theory "works," this is evidence that all its parts are true, including the mathematical parts. The analogy is questionable. Perhaps the parts of a typical computer program are related in a special way; maybe none of the parts are *idle*, in that each must play its role if inputs are to yield the right outputs. In contrast, it is well known that theories can be constructed that contain idle postulates. Just conjoin a routine scientific theory with the statement "undetectable angels exist." The predictive success of the conjunction is hardly evidence that the second conjunct is true. It also is well known that a theory can be empirically successful even when it is known to contain idealizations; its success in that case does not provide evidence that each of its parts is true. These points underscore the fact that Colyvan's analogy is only an analogy. It needs to be cashed out via a principle of confirmation. There is ample reason to grant that the special consequence condition is false, and the fact that the consequence of interest happens to be a proposition of pure mathematics does not automatically rescue the principle.

(p.212) VII. Three Kinds of Naturalism

I have described naturalism as a metaphysical doctrine, but an alternative formulation, methodological naturalism, has also been much discussed in connection with the conflict between evolutionary theory and creationism.[50] I want to add a third version of naturalism to this familiar pair. Here are all three:

> *Metaphysical Naturalism*: The only entities that exist are in nature.
>
> *Methodological Naturalism*: The only entities that scientific theories should postulate are in nature.
>
> *Evidential Naturalism*: The only entities whose existence we have evidence for are in nature.

Evolutionists who criticize creationism often emphasize that methodological naturalism doesn't entail metaphysical naturalism and that science is committed only to the former. I agree that the one doesn't entail the other, but I disbelieve them both. In fact, I think that all three naturalisms are flawed.

Before describing why I reject methodological naturalism, I want to note what this principle does not exclude. It does rule out mention of supernatural deities, but it does not proscribe deities that are not outside of nature. The Greek gods are okay, and so is Spinoza's God, who is identical with nature. If the point of this doctrine is to rule out deities in general, a stronger principle is needed.

Methodological naturalism is often defended by appeal to the thesis that claims about supernatural entities cannot be tested.[51] When testability is understood in the usual way, this is an overstatement. Consider, for example, the claim that:

> A supernatural intelligent designer caused everything in nature to be purple.

This statement about a supernatural being *is* testable. It is refuted by an observation statement—that my car isn't purple. A more promising way to develop the idea about testability is to think of testing contrastively. Consider, for example, the hypothesis that:

> (ID) A supernatural intelligent designer created the vertebrate eye.

Suppose we want to test this ID hypothesis against the hypothesis that:

> (NS) Natural selection created the vertebrate eye.

(p.213) *Both* hypotheses entail that vertebrates have eyes, and so both are testable in the usual, non-contrastive, sense of that term. But notice that the fact that vertebrates have eyes does not *discriminate* between ID and NS. To solve this discrimination problem, we need a set of observations that renders the two hypotheses unequally likely. So let's consider the observation that the vertebrate eye has features *F*. Is this outcome more probable if ID is true, or if NS is true? The problem is that we can't assess what probability the ID hypothesis confers on this observation unless we have independently justified information about the goals and abilities the putative designer would have if there were such a being. And we have no such information.[52]

If ID can't be tested against NS, it is equally true that NS can't be tested against ID. What's poor evolutionary theory to do then, if creationism isn't able to put up its dukes? The answer is that evolutionary biology allows one to formulate multiple hypotheses about the vertebrate eye and to test these against each other. One alternative to natural selection is drift. Talk of testing evolutionary theory is a bit like talk of testing chemistry. Evolutionary biology is a *field*. There are many possible evolutionary explanations of a particular trait's presence in this or that group of organisms, and biologists test these hypotheses against each other.

Just as methodological naturalism doesn't entail the metaphysical variety, it also is true that metaphysical naturalism does not entail the methodological version. This independence thesis seems to fly in the face of a simple argument. If there are no supernatural entities, then scientific theories that postulate such things are false. And scientific theories shouldn't include falsehoods. QED.

One flaw in this line of reasoning becomes clear when we reflect on the fact that scientists aren't making a mistake when they formulate and test theories that contain idealizations.[53] Physicists construct models that postulate point particles, spherically symmetrical planets, and frictionless planes; evolutionary biologists construct models that postulate infinitely large populations. Why do scientists bother to do this if they know in advance that such theories are false? One justification comes from the fact that idealizations are often known to be *harmless*: they don't much affect the predictions that the theory makes.[54] If *I* is a harmless idealization of a more realistic theory *T*, then you can see how accurately *T* predicts by seeing how accurately *I* predicts. And if you think that *T* is true, you might as well use *I* to **(p.214)** make your predictions. None of this should give pause to scientific realists.[55]

If metaphysical naturalism is true, then there are no numbers, at least not as Platonistically conceived. But this doesn't mean that scientists should abstain from formulating and testing hypotheses that entail that numbers exist. Science *needs* to postulate numbers, whether or not such things really exist. So even if we *knew* that metaphysical naturalism is true, that would not oblige us to play by the rules of methodological naturalism. This line of reasoning requires that we recognize that there is an ambiguity in talk about “the things that science needs to postulate.” There are the things that scientists are entitled to *believe* exist. And there are the things that can be postulated in the models that scientists *use* without fully believing them. A useful fiction known to be such can't be postulated in the first sense, but it obviously can be in the second. So metaphysical naturalism does not entail methodological naturalism, where the latter prohibits formulating and testing theories that entail the existence of numbers, and using some of them to make predictions. And if a theory can be explanatory without being 100 per cent true, then theories with false metaphysics can even be explanatory.

Although the legitimate use of useful fictions is enough to show that metaphysical naturalism doesn't entail methodological naturalism, it is important not to exaggerate the similarity between idealizations in science and science's use of mathematics. If Platonism is false, then frictionless planes and numbers have in common their nonexistence. But still, there is a world of difference. For one thing, numbers aren't just *useful*; they are *indispensable*. As Friedman notes, we use mathematical language in describing our observations, and theories that hope to do justice to those observations must quantify over numbers. In contrast, most idealizations in the sciences are, in an important sense, dispensable. Physicists can write models in which balls experience friction when they roll down incline planes and geneticists can write models in which finite population size is taken into account. These models are more complicated, but they are not beyond our ken. Maddy points out that idealized models are often more mathematically tractable,[56] but that does not mean we are hopelessly in the dark without these idealizations. Another difference between frictionless planes and **(p.215)** numbers is that we have empirical evidence that planes aren't perfectly frictionless. However, we have no empirical evidence that there are no numbers.[57]

There are two senses of indispensability. Electrons are indispensable in the electron theory if they aren't *idle postulates* within that theory—if, in using the theory to predict and explain, we need to use the hypothesis that electrons exist and have the various properties that the theory describes. However, electrons are not indispensable *in theories that compete with the electron theory*. This is why we can *test* whether electrons exist. Numbers are different. They are indispensable in the electron theory, and they also are indispensable in theories that compete with the electron theory by making different quantitative predictions. Electrons are indispensable *within* a theory, but numbers are indispensable *across* theories. This is why evidence for the electron theory can provide evidence for the existence of electrons, but not for the existence of numbers. Indispensability across theories is just what *ruins* our chances of providing empirical evidence in favor of a postulate.

What about the third naturalism described, evidential naturalism? Is this the kernel of truth that remains after we recognize that methodological naturalism is an overstatement? The problem here concerns the fact, noted earlier, that there can be empirical tests of statements in pure mathematics, including existence claims. Benacceraf's question is—how can what we see *in* nature be evidence for the existence of things that are *outside* of nature? He thinks the answer to this question is that natural observations can't have this evidential import. But why not? How can what we see *in New Jersey* be evidence for the existence of things that are outside of New Jersey? Of course they can be. I grant that the puzzle about Platonism is more difficult than the puzzle about New Jersey. But there is a useful parallelism; observing one thing can be evidence for the existence of another if the two are “connected.” This need not involve a causal relationship, although that is one possibility. Maybe propositions in the calculus of variations aren't *causally* connected to the soap bubbles that Plateau observed, but, evidently, they are connected.[58]

Even if evidential naturalism is false, there is a more limited evidential thesis that is correct. This is exemplified by the models FIS and FDS. Though each entails the existence of numbers, each is an empirical hypothesis and they can be tested against each other. The Platonistic commitments are what they have in common; testing them against each other does not **(p.216)** test that shared presupposition. When scientific tests are organized in this way, the observations do not provide evidence for or against the existence of supernatural entities. However, the same can be said about the irrelevance of scientific tests to the question of whether physical objects exist. The contrastive thesis is right, but it isn't distinctively *naturalistic*.

VIII. Concluding Comments

I started this paper by quoting Jacques Monod's statement that “any confusion between the ideas *suggested* by science and science itself must be carefully avoided.” Monod's dictum is relevant to atheistic evolutionists and to creationists alike. Evolutionary theory apparently “suggests” to many people that there is no God, but the theory has no such implication. There is another confusion that needs to be avoided as well. We must be careful not to confuse the implications of a well-confirmed theory for which we have empirical evidence and the implications for which we have none. If Hempel's special consequence condition were true, there would be no such distinction. But the special consequence condition is wrong.

Although evolutionary theory is silent on the question of whether there is a God, it has implications about supernatural entities of another kind. The theory quantifies over numbers just as much as it quantifies over genes, organisms, and populations. If numbers are Platonistic entities, evolutionary theory violates the requirements of methodological naturalism. This does not mean that evolutionary biologists are obliged to be mathematical Platonists. They have the option of adopting a Carnapian attitude,[59] granting that their theory uses the framework of numbers while asserting that they adopt this framework for its convenience, not because they can produce a good argument for its truth. The theory of evolution, like other empirical theories, is a whole with many parts and its parts differ in their standing. Some are well attested empirically, others are known to be useful fictions, and still others are indispensable, not just useful. Duhem and Quine defended epistemological holism, arguing that it is whole theories that are tested, not their parts.[60] But the fact of the matter is that theories can often be pulled apart and their parts tested separately.[61] Scientists don't **(p.217)** take the same attitude to all the parts of their theories, nor are they obliged to do so.[62] Putnam says that anti-realists are guilty of intellectual dishonesty, since they daily presuppose the existence of the entities that they say they forswear.[63] My response is that there is nothing dishonest about saying that you are using a model that contains assumptions you are unable to test, nor is it dishonest to admit that some of the entities postulated by the model are useful fictions. What is dishonest is pretending that you have evidence for all the propositions on a list when you know that you have evidence only for some.

Although the two questions about metaphysics that I have discussed concern supernatural beings, God and numbers, the antiholistic view of evidence that I have described also applies to metaphysical questions about nature itself. Evolutionary theory entails that there are physical objects that are external to my mind; it therefore entails that solipsism is false. However, this does not mean that the evidence we have for evolutionary theory is evidence against solipsism. The models that evolutionary biologists test against each other all *assume* that solipsism is false. This proposition is assumed, not tested, just like the framework of numbers.[64]

Although evidence is the epistemological concept on which I have concentrated, it is worth considering how a nonholistic view of evidence bears on the role of “acceptance” in the scientific enterprise. Does rational acceptance obey a closure principle? If you (rationally) accept a hypothesis H and you know that H entails C, must you also (rationally) accept C? My claims about evidence do not entail that this principle is false. I think that evolutionary biology provides lots of evidence for evolutionary theory but none for the existence of numbers. If the theory entails that numbers exist, are biologists being irrational if they endorse the theory while being agnostic about the Platonism? Of course, most scientists have not considered the fact that scientific theories entail that numbers exist. But let us bring some of them into a classroom and make them listen. There is no reason why biologists should not avail themselves of the simple response that they take no stance on $M\&E$, but wish to affirm merely that “if M, then E.” Epistemologists can have their closure principle and biologists can have the parts of their theory that they care about.[65]

(p.218) There have been many versions of “the theory of evolution;” some have been more metaphysical while others have been more austere. The present inquiry does not require that one version must be the “real” theory while all the others are inflated or deflated pretenders. David Hull argues that theories are historical entities, like biological species; theories are not defined by their including this or that essential proposition, but by their historical relationships.[66] Hull's main example is Darwinism. For Hull, what makes people Darwinians is their relationship to others, not their assent to a list of required beliefs. I have argued that “evolutionary theory” is silent on the question of whether there is a God, but that it entails the existence of numbers. This pair of claims does not require an essentialist view of what evolutionary theory is. Whether atheistic evolutionism is the “real” theory of evolution does not matter. The point is that evolutionary biology provides no evidence that there is no God. The same point holds of theistic evolutionism; whether or not this is the “real” evolutionary theory, it includes an element for which evolutionary biology provides no evidence. Similar remarks pertain to the question of whether numbers exist. Mathematical evolutionary theory entails that there are numbers, but nonmathematical propositions about evolution (e.g., that humans and chimps have a common ancestor) do not. Without having to decide which of these is the “real” theory of evolution, we can simply note that there is no biological evidence that numbers exist. So what is the difference between God and numbers? You need numbers to do science. But that isn't evidence that numbers exist and that God does not.

References

Bibliography references:

Benacceraf, P. (1973), “Mathematical Truth,” *Journal of Philosophy* 70: 661–79.

Brandon, R. and Carson, S. (1996), "The Indeterministic Character of Evolutionary Theory—No 'No Hidden Variables Proof' but No Room for Determinism Either," *Philosophy of Science* 63: 315–37.

Burnham, K. and Anderson, D. (1998, 2002), *Model Selection and Inference—a Practical Information-Theoretic Approach* (New York: Springer).

Carnap, R. (1956), "Empiricism, Semantics, and Ontology," *Meaning and Necessity* (Chicago, IL: University of Chicago Press).

Clifford, W. K. [1877] (1999), "The Ethics of Belief," *The Ethics of Belief and Other Essays* (Amherst, NY: Prometheus Books).

(p.219) Colyvan, M. (2001), *The Indispensability of Mathematics* (New York: Oxford University Press).

—— (2004), "Indispensability Arguments in the Philosophy of Mathematics," *The Stanford Encyclopedia of Philosophy,* E. Zalta (ed.), ⟨http://plato.stanford.edu/entries/mathphil-indis/⟩.

Courant, R. and Robbins, H. (1941), *What is Mathematics?* (London: Oxford University Press).

Crow, J. and Kimura, M. (1970), *An Introduction to Population Genetics Theory* (Minneapolis, MN: Burgess Publishing Co.).

Darwin, C. (1903), *More Letters of Charles Darwin*, ed. F. Darwin and A. Seward (London: John Murray).

Dembski, W. (1998), "Introduction," in W. Dembski (ed.), *Mere Creation—Science, Faith, and Intelligent Design* (Downers Grove, IL: InterVarsity).

Dennett, D. (1995), *Darwin's Dangerous Idea* (New York: Touchstone).

Diaconis, P. (1998), "A Place for Philosophy? The Rise of Modeling in Statistical Science," *Quarterly of Applied Mathematics* 56: 797–805.

Dretske, F. (1970), "Epistemic Operators," *Journal of Philosophy* 67: 1007–23

—— (1972), "Contrastive Statements," *Philosophical Review* 81: 411–37.

Duhem, P. [1914] (1954), *The Aim and Structure of Physical Theory* (Princeton, NJ: Princeton University Press).

Earman, J. (1986), *A Primer on Determinism* (Dordrecht: Reidel).

Easwaran, K. (2008), "Probabilistic Proofs and Transferability," unpublished.

Edwards, A. (1972), *Likelihood* (Cambridge: Cambridge University Press).

Everett, T. (2006), “Antiskeptical Conditionals,” *Philosophy and Phenomenological Research* 73: 505–36.

Field, H. (1980), *Science without Numbers* (Oxford: Blackwell).

—— (1989), *Realism, Mathematics, and Modality* (Oxford: Blackwell).

Friedman, M. (2001), *Dynamics of Reason* (Stanford, CA: CSLI Publications).

Hacking, I. (1965), *The Logic of Statistical Inference* (Cambridge: Cambridge University Press).

Hellman, G. (1999), “Some Ins and Outs of Indispensability—a Modal-Structural Approach,” in A. Cantini, E. Casari, and P. Minari (eds.), *Logic and Foundations of Mathematics* (Dordrecht: Kluwer), 25–39.

Hempel, C. (1965), “Studies in the Logic of Confirmation,” *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (New York: Free Press).

Hull, D. (1988), *Science as a Process—An Evolutionary Account of the Social and Conceptual Development of Science* (Chicago, IL: University of Chicago Press).

Jablonka, E. and Lamb, M. (2005), *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life* (Cambridge, MA: The MIT Press).

James, W. (1897), “The Will to Believe,” in *The Will to Believe and Other Essays in Popular Philosophy* (New York: Longmans Green & Co.), 1–31.

(p.220) Keller, J. (1986), “The Probability of Heads,” *American Mathematical Monthly* 93: 191–97.

Kotzen, M. (2007), “Dragging and Confirming,” unpublished.

Kyburg, H. (1970), “Conjunctivitis,” in M. Swain (ed.), *Induction, Acceptance and Rational Belief* (Boston: Reidel), 55–82.

Lenski, R. and Mittler, J. (1993), “The Directed Mutation Controversy and Neo-Darwinism,” *Science* 259: 188–94.

Lyon, A. and Colyvan, M. (2008), “The Explanatory Power of Phase Spaces,” *Philosophia Mathematica* 16: 227–43.

Maddy, P. (1992), “Indispensability and Practice,” *Journal of Philosophy* 89: 275–89.

—— (1993): “Taking Naturalism Seriously,” in D. Prawitz, B. Skyrms, and D. Westgerstahl (eds.), *Proceedings of the Ninth International Congress of Logic, Methodology, and Philosophy of Science* (Amsterdam: North Holland Publishing Co.).

—— (1997), *Naturalism in Mathematics* (Oxford: Clarendon Press).

Mayr, E. (1961), “Cause and Effect in Biology,” *Science* 134: 1501–6.

Monod, J. (1971), *Chance and Necessity* (New York: Alfred Knopf).

Numbers, R. (2003), “Science Without God—Natural Law and Christian Beliefs: An Essay on Methodological Naturalism,” in D. Lindberg and R. Numbers (eds.), *When Science and Christianity Meet* (Chicago, IL: University of Chicago Press), 265–86.

Papineau, D. (1993), *Philosophical Naturalism* (Oxford: Blackwell).

Parsons, C. (1980), “Mathematical Intuition,” *Proceedings of the Aristotelian Society* 80: 145–68.

Pennock, R. (1999), *Tower of Babel—The Evidence against the New Creationism* (Cambridge, MA: The MIT Press).

Pigliucci, M. (2001), *Phenotypic Plasticity—Beyond Nature and Nurture* (Baltimore, MD: Johns Hopkins University Press).

Plantinga, A. (2000), *Warranted Christian Belief* (Oxford: Oxford University Press).

—— (2009), “Science and Religion—Why Does the Debate Continue?,” in M. Stewart (ed.), *Science and Religion in Dialogue* (Oxford: Blackwell, 2010), 301–17.

Popper, K. (1959), *Logic of Scientific Discovery* (London: Hutchinson).

Provine, W. (1989), “Progress in Evolution and Meaning in Life,” in M. Nitecki (ed.), *Evolutionary Progress* (Chicago, IL: University of Chicago Press).

Putnam, H. (1971), *Philosophy of Logic* (Boston, MA: Allen and Unwin).

Quine, W. (1953), “Two Dogmas of Empiricism,” *From a Logical Point of View* (Cambridge, MA: Harvard University Press), 20–46.

Railton, P. (1978), “A Deductive-Nomological Model of Probabilistic Explanation,” *Philosophy of Science* 45: 206–26.

(p.221) Reichenbach, H. (1938), *Experience and Prediction* (Chicago, IL: University of Chicago Press).

Resnik, M. (1995), “Scientific vs. Mathematical Realism—The Indispensability Argument,” *Philosophia Mathematica* 3: 166–74.

—— (1997), *Mathematics as a Science of Patterns* (Oxford: Oxford University Press).

Rosenberg, A. (1994), *Instrumental Biology or the Disunity of Science* (Chicago, IL: University of Chicago Press).

Royall, R. (1997), *Statistical Evidence—A Likelihood Paradigm* (Boca Raton, FL: Chapman and Hall).

Ruse, M. (2000), *Can a Darwinian be a Christian?* (Cambridge: Cambridge University Press).

Sakamoto, Y., Ishiguro, M., and Kitagawa, G. (1986): *Akaike Information Criterion Statistics* (New York: Springer).

Sarkar, Sahotra (2005), “Neo-Darwinism and the Problem of Directed Mutations,” *Molecular Models of Life: Philosophical Papers on Molecular Biology* (Cambridge, MA: The MIT Press), 287–301.

Sober, E. (1993a), “Mathematics and Indispensability,” *Philosophical Review* 102: 35–58.

—— (1993b), *Philosophy of Biology* (Boulder, CO: Westview Press).

—— (1999), “Physicalism from a Probabilistic Point of View,” *Philosophical Studies* 95: 135–74.

—— (2004), “Likelihood, Model Selection, and the Duhem-Quine Problem,” *Journal of Philosophy* 101: 1–22.

—— (2008), *Evidence and Evolution—The Logic Behind the Science* (Cambridge: Cambridge University Press).

—— (2009), “The Birds Outside—Reichenbach's Cubical Universe and the Problem of the External World,” *Synthese*, forthcoming.

—— (2010), “Evolutionary Theory and the Reality of Macro-Probabilities,” in E. Eells and J. Fetzer (ed.), *Probability in Science* (LaSalle, IL: Open Court).

Swinburne. R. (2004), *The Existence of God* (Oxford: Oxford University Press).

Tymoczlo, T. (1979), “The Four-Color Problem and its Philosophical Significance,” *Journal of Philosophy* 76: 57–83.

Notes:

(1) My thanks to Juan Comesaña, Mark Colyvan, James Crow, Hartry Field, Geoffrey Hellman, Richard Lewontin, Bence Nanay, Peter Nichols, Alexander Paseau, Alexander Rosenberg, Michael Ruse, Carolina Sartorio, Larry Shapiro, Alan Sidelle, and Michael Titelbaum for useful comments.

(2) Hempel (1965: 31).

(3) This illustrates Dretske's (1970) thesis that the evidence relation involves a non-penetrating epistemic operator.

(4) I am indebted to Kotzen (2007) for this way of representing the point.

(5) I argue in chapter 2 of Sober (2008) that the argument from design is defective for reasons that are independent of whether the theory of evolution is true.

(6) Plantinga (2000).

(7) It is not in dispute that evolutionary theory conflicts with certain forms of Biblical literalism; in this respect, the theory is no different from theories in other sciences.

(8) In Sober (1999) I defend a screening-off formulation of the thesis that physics is causally complete and discuss its relationship to the thesis of supervenience and also to the thesis that mental properties are causally efficacious.

(9) Lenski and Mittler (1993: 188).

(10) Consider the simplest genetic case—a site in an organism's genome has nuceleotide G, and we are considering the three mutations it might experience—to A, to C, and to T. There are six possible environments we might consider that are distinguished from each other by the ordering of the fitnesses of these three mutations. To check for mirroring, we'd need to consider eighteen fitnesses and eighteen mutation probabilities. The concept of perfect mirroring is defined as it was before, but we'd like also to have a concept of less than perfect mirroring as well. I won't address the question of how this should be understood.

(11) I'm here defining what it means for mutations to be directed towards what is good for organisms. One also could consider other, less benign, forms of directedness. For a more detailed discussion of what directed mutation means, see Jablonka and Lamb (2005) and Sarkar (2005).

(12) I say “prima facie,” because it may be that there are negatives associated with such a mechanism (e.g., if it imposes a huge energetic cost on the organism) so that, on balance, it would not be selectively favored over alternatives.

(13) There are theoretical dimensions to the question of directed mutation that are worth considering. For example, many organisms are capable of *phenotypic plasticity* (Pigliucci 2001). An iguana can change its color without needing to mutate. In what circumstances will directed mutation provide advantages that phenotypic plasticity cannot? One difference might be that mutation in an adult organism often leaves its own phenotype unchanged while altering the phenotype of its offspring, whereas phenotypic plasticity changes the adult's own phenotype. An iguana that lives in a highly changeable environment may do better with the latter than it would do with the former.

(14) Sober (2010). Compare Brandon and Carson (1996), who argue that evolutionary biology shows that there are no such hidden variables.

(15) Crow and Kimura (1970: 87).

(16) Plantinga (2009) distinguishes “Darwinism” from “unguided Darwinism” and interprets the latter as incompatible with the existence of God. For Plantinga, Darwinism is a scientific theory whereas unguided Darwinism is that theory with a “metaphysical or theological add-on.” I agree that the distinction is important, but think that the notion of unguidedness that is central to evolutionary biology is itself theologically neutral.

(17) Darwin (1903, 1: 192.

(18) Dembski (1998: 20–1).

(19) Swinburne (2004).

(20) Benacceraf (1973).

(21) Field (1980, 1989).

(22) See Papineau (1993) for a discussion of this issue in the philosophy of mathematics within the larger context of naturalism.

(23) Hacking (1965); Edwards (1972); Royall (1997).

(24) I here omit the second clause in the Law of Likelihood, which says that the degree of favoring is to be measured by the likelihood ratio. This additional claim won't be relevant in what follows.

(25) You don't have to be a Bayesian to find merit in the Law of Likelihood; “likelihoodists” hold that the principle describes how evidence should be interpreted even when you are unwilling to assign prior or posterior probabilities to the hypotheses under test.

(26) Dretske (1972) is the *locus classicus* for the argument that a number of philosophically interesting concepts are contrastive.

(27) Duhem (1914) and Quine (1953).

(28) Sakamoto *et al*. (1986: 84).

(29) AIC provides an estimate of a model's predictive accuracy, while the Law of Likelihood describes how one should decide whether the data favor the hypothesis that M_1 is true over the hypothesis that M_2 is true. Models known to be false can be good predictors; in fact, a model known to be false can sometimes be known to be a better predictor than a model known to be true. For further discussion, see Sober (2008).

(30) The contrastive character of likelihood comparisons and of AIC scoring throws light on Maddy's (1997: 156) thesis that “the success of a theory involving certain mathematical existence assumptions and corresponding physical structural assumptions is not regarded as confirming evidence for those assumptions.” She is right, not just in describing this aspect of scientific practice, but in endorsing it.

(31) Courant and Robbins (1941: 385–91).

(32) Tymoczlo (1979).

(33) Easwaran (2008).

(34) Friedman (2001: 35–6).

(35) Consider, for example, the question of whether evolutionary theory shows that taxa don't have essences.

(36) Quine (1953) and Putnam (1971). See Colyvan (2004) for a useful formulation of the argument and a survey of criticisms.

(37) Sober (1993b).

(38) Parsons (1980).

(39) Maddy (1992).

(40) Hellman (1999: 32).

(41) Ibid. (30).

(42) Kotzen (2007).

(43) Hellman (1999: 30).

(44) Ibid. (33).

(45) Colyvan (2001: 128).

(46) Sober (1993a).

(47) Resnik (1997: 46–7).

(48) Colyvan (2001, 2004); Hellman (1999); and Resnik (1995, 1997).

(49) Colyvan (2001: 134).

(50) Numbers (2003).

(51) Pennock (1999).

(52) Sober (2008: ch. 2).

(53) Maddy (1992).

(54) Sober (1993b).

(55) However, there is another reason to formulate and test hypotheses that are known to include idealizations—namely, that we want to find theories that make accurate predictions about observations and that all the theories we are able to assess include idealizations. Considering this instrumentalist suggestion would take us too far afield, however.

(56) Maddy (1992).

(57) Hellman (1999: 31).

(58) Colyvan (2001: ch.3) and Lyon and Colyvan (2008) argue that pure mathematics sometimes bears an explanatory, though not a causal, relation to physical phenomena.

(59) Carnap (1956).

(60) In Sober (1993a) I distinguish two forms of holism. Nondistributive holism holds that the confirmation of a whole theory never means that any of its parts have been confirmed. Distributive holism says that the confirmation of a theory entails that each of the parts is confirmed. The latter view is wrong if Hempel's special-consequence condition is wrong.

(61) Sober (2004).

(62) Maddy (1992).

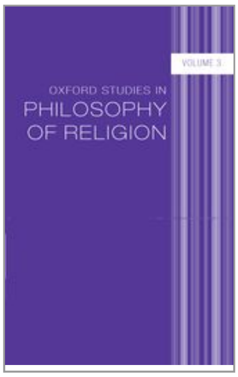
(63) Putnam (1971: 57).

(64) My arguments against the special-consequence condition do not mean that we have no empirical evidence that solipsism is false. Reichenbach (1938) argued that we do; I consider his argument in Sober (2009).

---

(65) See Everett (2006) for discussion of “antiskeptical conditionals” in connection with skepticism about knowledge.

(66) Hull (1988).



Oxford Studies in Philosophy of Religion Volume 3

Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

Geachianism

Patrick Todd

DOI:10.1093/acprof:oso/9780199603213.003.0011

Abstract and Keywords

In his 1977 monograph, Providence and Evil, P.T. Geach articulated a position on divine omniscience and future contingents that has been widely interpreted as a version of (or precursor to) ‘open theism’. In short, Geach argued that the future is (in a particular way) mutable. Unfortunately, however, and despite Geach's philosophical eminence, the distinctive view which he proposed has had little impact on the massive literature on such topics. Admittedly, this is perhaps because Geach did not systematically develop the view he was proposing. The goal of this paper is thus to further develop Geach's view. Though the resulting view–Geachianism–is certain to be controversial, the chapter argues that it deserves the status of a theoretical contender in these debates.

Keywords: P.T. Geach, Geachianism, divine foreknowledge, the mutability of the future, open theism, omniscience

In his 1977 monograph, *Providence and Evil*, P. T. Geach articulated a position on divine omniscience, future contingents, and providence that has been widely interpreted as a version of (or precursor to) "open theism." Unfortunately, however, and despite Geach's philosophical eminence, the distinctive view that he proposed has had little impact on the massive literature on such topics.[1] Indeed, those philosophers—with one primary exception—who have read *Providence and Evil* have either failed to notice that Geach was arguing for a radically new view on the logic of future contingents or have not thought it worth commenting on. Admittedly, this is perhaps because Geach himself did not systematically develop the distinctive view he was proposing. I aim to do so here, albeit in a non-exhaustive, preliminary way. I call the resulting view "Geachianism."[2] Though Geach's view is certain to be controversial, I claim that it deserves the status of a theoretical contender in these debates. Moreover, a discussion of Geach's view illuminates various subtle and fascinating points about the dilemma of freedom and foreknowledge not previously noticed in the literature on these topics.

The plan of the paper is as follows. To set the stage, I first briefly present the most sophisticated characterization of "open theism" on offer, one recently developed by Alan Rhoda.[3] Having other "open-theistic" views on the table will allow us better to understand and situate Geach's view. Second, **(p.223)** I explain and motivate Geach's view of future contingents, and respond to some of the objections it faces, notably those made by Jonathan Kvanvig, in what is (as far as I know) the only explicit discussion of Geach's view in the philosophical literature.[4] Next, I show how such a view enables a particular reply to the argument for the incompatibility of foreknowledge and freedom of Nelson Pike's seminal 1965 paper, "Divine Omniscience and Voluntary Action." Historically, however, open theism has not been a mere reaction to this argument, but has had distinctively theological motivations stemming from a (putative) "plain reading" of certain scriptural passages and the desire for a particular account of divine providence. An application of Geach's view to such topics will show that it has certain advantages over traditional "open" views.

I. Rhoda on "Generic Open Theism"

First, Rhoda thinks all open theists must be committed to theism (rather than *process theism*) and indeterminism, or the “causal openness” of the future.[5] As Rhoda notes, the primary, historical motivation for the adoption of indeterminism amongst theists has been the (putative) existence of *libertarian freedom*, the exercise of which requires indeterminism. Now, on Rhoda's view, open theists depart from other indeterminist theists in maintaining that if it is causally open whether some state of affairs occurs at some time, then it is impossible for God to know beforehand that it will or will not occur. Ockhamists and Molinists, for instance, maintain that, of the causally possible futures, God knows that some *particular* future is the actual one. Open theists (says Rhoda) demur; on open theism, the future is “epistemically open” for God in the sense that there are some states of affairs such that God neither knows that they *will* obtain nor that they *will not* obtain. This is the central feature of Rhoda's characterization of “generic open theism.”

In brief, the characterization is “generic” in the sense that it is neutral between the camps of the two most prominent open-theist positions. According to one camp, future contingents—propositions which say of some undetermined event that it will happen—are never true. On this view, **(p.224)** if it is causally open whether X obtains at *t3*, then it cannot even be *true* at (some earlier time) *t*1 that X will obtain at *t3*. And if it is not true at *t*1 that X will obtain at *t*3, then God (given his essential omniscience) does not believe at *t*1 that X will (or will not) obtain at *t*3. The future is thus epistemically open for God in the sense at stake. We may call those in this camp “open- future open theists.” In this camp, there is yet another important division: those who maintain that the relevant future contingent propositions are neither true nor false (hence denying bivalence) and those who believe such propositions are uniformly false. This is an important (and often neglected) distinction, and interested readers would do well to consult Rhoda's work on these issues.[6] Prominent open-future open theists include (anachronistically, controversially, and somewhat tendentiously) Aristotle[7] and J. R. Lucas (who deny bivalence),[8] and Charles Hartshorne[9] and Rhoda himself (who do not).[10]

The other prominent open-theist camp does not maintain that future contingent propositions are not *true*, but that it is impossible for God to know them, despite their being true. This impossibility generates the relevant divine epistemic openness. According to this version of open theism, it is logically impossible for God to know how future indeterminacies (such as our free decisions) will unfold. Assuming indeterminism, then it follows that the future is epistemically open for God. Taking after Rhoda, we may call this version “limited foreknowledge open theism.” Richard Swinburne **(p.225)** and William Hasker have been the primary defenders of this version of open theism, though Peter van Inwagen has recently defended it as well.[11]

Rhoda's analysis of “generic open theism” helpfully shows what these versions of open theism have in common. Both maintain that the future is epistemically open for God. That is, both maintain that there are some possibly future states of affairs such that God knows neither that they will nor will not obtain. And it would seem that this thesis is essential to open theism; it apparently captures its defining feature—what puts the “open” in “open theism,” as Rhoda says. Indeed, how could one be an open theist if one *denies* that there are some things about which God neither knows that they will nor will not happen? If for every state of affairs, God knows that it will happen or that it will not happen, the future would plainly be epistemically *settled* for God, as Rhoda puts it. Of the myriad causally possible futures, God would know that some one of them is *the actual one*. If we think of causal possibility in terms of branches off a single path, God would see that one such branch is “lit up” with what Nuel Belnap and Mitchell Green have called “the thin red line,” indicating that *it* is the special branch—the one that uniquely *will* obtain, the one which is such that it is *going to happen*.[12] And this is just the traditional view of divine foreknowledge. Such considerations would appear to be decisive, but they are decisive only under the supposition that something that is going to happen cannot later be such that it is no longer going to happen—only under the supposition, that is, that *the future cannot change*. Geach denies this. To these issues I now turn.

II. Geach and the Mutability of the Future

As I remarked at the outset, it would be an understatement to say that Geach's views about future contingents have not been widely influential. But what is the view?

To begin, what is distinctive of Geach's view is that it is possible for something to be such that, at *t*1, it will happen at *t*3, but at *t*2 such that it will not happen at *t*3. That is, at *t*1, it was true that X would happen at *t*3, but at *t*2, something intervened to make it the case that X would not happen at *t*3. The future thus *changes* in the sense that something was going to happen, but now no longer is going to happen, and does not happen. In support of his case, Geach points to the logic of *prevention*. According to Geach, those things are *prevented* from happening which were such that they were **(p.226)** *going to happen* but nevertheless do not happen. Moreover, if something is *going to happen*, then it *will* happen; these expressions are logically equivalent.[13] The result is that what will occur as of one time may later be such that it will not occur. It seems best to let Geach (largely) speak for himself. In considering one example, Geach asks rhetorically:

> But what then is prevented? Not what did happen, but assuredly what was going to happen. The aeroplane was going to crash into the sea and 100 men were going to be drowned; the pilot's prompt action prevented this. For not everything that does not happen is prevented: only what was going to happen.[14]

Geach says he takes it as a “truism” that “anything that is prevented is something that was going to happen but didn't happen.” He then asks:

> But if something did happen, doesn't this show it was after all going to happen? Certainly; but not that it *always* was going to happen. Perhaps, before the preventative action was taken, not this but something else was going to happen; but then the preventative action was taken, and after that *this* was going to happen and did happen. Before the pilot's daring manoeuvre, the plane was going to crash; but after that the plane was going to land safely and did land safely.[15]

Geach goes on:

> So what was going to happen at an earlier time may not be going to happen at a later time, because of some action taken in the interim. This is the way we can change the future: we can and often do bring it about that it will not be the case that *p*, although before our action it was going to be the case that *p*; it was right to say, then, ‘It is going to be the case that p.’ Before the operation it was right to say ‘Johnny is going to bleed to death from the injury’: after the operation this was no longer the case.[16]

So the view is that the future is *mutable* in a particular way. On open-future views, we often so act that we give the future a more determinate shape; we sometimes make it so that (previously) what only *might* or *will probably* happen becomes such that now it simply *will* happen. Geach's view is different. We change the future in a more radical way: we make it so that what will happen ends up not happening, or that what previously was such that it will not happen *does* happen. What will happen changes. No other view—“open” or otherwise—maintains this distinctive thesis.

I believe we can best understand Geach's position by contrasting it with the other possible views on whether a future contingent proposition can **(p.227)** change its truth-value over time. For simplicity, I consider only other *tensed* views which take it that in saying “X will happen at *t*10,” one expresses a proposition in which the future tense is fundamental and irreducible; one does *not* express a proposition like “X happens at t10.” Consider:

> (P) X will happen at *t*10.

Suppose (P) is future contingent. Can (P) change its truth-value over time? There are three (somewhat arbitrarily named) options we must consider. Intuitively, the three options are no change, change in one direction only, and change in both directions. Consider first:

> **The standard tensed theory of time**: (P) cannot go from being false to being true, and cannot go from being true to being false until *t*10 is present.

According to what I am calling the "standard" tensed view, (P) cannot go from being false to true, and can only go from true to false *once t10 is present or in the past*. On the tensed view, recall, a proposition like (P) is *not* reducible to a tenseless proposition; the tense is fundamental or basic. So while "X will happen at *t*10," if true, has always been true up until *t*10, it becomes false at and after *t*10. That is, if *t*10 is in the present or past, it is false that something *will* happen at *t*10.

The next option to consider is the one endorsed by those who believe there are no true future contingent propositions, such as open-future open theists. If future contingents are uniformly false, then we have:

> **The open-bivalentist tensed theory of time**: (P) can go from being false to being true, but cannot go from being true to being false until *t*10 is present.

And if future contingents are neither true nor false:

> **The open-non-bivalentist tensed theory of time**: (P) can go from being neither true nor false to being true, but cannot go from being true to being neither true nor false (or false), until *t*10 is present.

On the view in question, it is not true that X *will* happen at *t*10 unless it is causally necessary that X happens at *t*10. Such causal factors may not be in place at *t*1, but may come into place at *t*3. Thus, at *t*3 it becomes true (whereas it was previously false, or neither true nor false) that X will happen at *t*10. In this way, a proposition about the future can become true. Once X is determined to occur at *t*10, it cannot later become undetermined that it occurs at *t*10. Thus, (P) cannot go from being true to being false (or neither true nor false), at least until *t*10 is present or past, in which case (P) will be false for the reasons just given on the "standard" view.

(p.228) The last view is Geach's:

> **Geachianism**: (P) can go from being true to being false, and from being false to being true, but must remain false when *t*10 is present or past.

We have seen the reasons Geach cites for thinking that (P) can change its truth values in these ways. X could have been going to happen, but got prevented from happening. In that case, (P) goes from true to false. On the other hand, perhaps X was not going to happen, but something arose that prevented what was going to happen, with the result that now X *is* going to happen. In that case, (P) goes from being false to being true. In sum, then, we have three intuitive options concerning whether (P) can change in truth value (before *t*10): no change in truth value (the standard view), change in one direction only (open-future views), and change in both directions (Geachianism).

III. On Behalf of Geachianism

Well, what can be said on behalf of Geachianism? The first thing to say on its behalf is simply that we very often *say* that things were going to happen but were prevented from happening. The plane was going to crash, but it didn't. Johnny was going to bleed to death, but he didn't. And not only do we say these things, but they seem true. Once we recognize it as such, the mutability of the future actually turns out to be a presupposition of a wide range of our discourse. So the first thing to say on behalf of Geachianism is that it very often seems true to say that the future changed.

Of course, many will want to attempt to explain away the appearances here. The second thing to say on behalf of Geachianism is that explaining it all away turns out to be surprisingly difficult. Recall Geach's example of the plane crash; the plane was going to crash, but the pilot prevented this from happening. The most obvious way to try to explain away the appearance of a changing future here is to maintain that it was never true that the plane was going to crash *simpliciter*, but only that it was going to crash *unless some preventative action is taken*. In other words, one interprets “The plane is going to crash” as elliptical for a merely *conditional* claim. Geach anticipates this response. As he says:

> I am prepared for the objection that I have been systematically equivocating upon two senses of “going to happen”: what actually will happen, and—well, **(p.229)** what? What will happen if nothing prevents it, perhaps. My complaint now is not that this phrase just boils down to “What will happen unless it doesn't”; for I myself do not identify what does not happen with what is prevented. But I do say that the explanation is useless. For what is prevented is always something that is going to happen, in the very sense of “going to happen” that we are supposed to be explaining; “prevent” has to be explained in terms of *this* “going to happen,” so we cannot use “prevent” to explain it. As for the “actually” in “what will actually happen,” it has no more logical force than a thump on the table has.[17]

In other words, Geach claims that some notion of something's “going to happen” is required for an analysis of prevention. The point is subtle, yet important. What is prevented is that which was *in some sense* going to happen. But in what sense? Here, one cannot say: what was going to happen unless something prevented it. For this sense of “going to happen” already *includes* the notion of prevention, and what we are looking for is the sense in which what is prevented was “going to happen.” According to Geach, prevention cannot be properly analyzed without appeal to what is “going to happen.” Thus, this way of explaining away the appearance of a changing future fails.

Of course, this strategy of responding to Geach would be back on the table if one could provide an alternative account of prevention—an account that does not identify the prevented with what was going to happen but did not. Again, the objector wishes to explain away the appearance of a changing future in Geach's examples; it was never really true that the plane was going to crash, only that it was going to crash unless prevented from crashing. But here Geach complains that what is prevented *just is* what was going to happen. Geach's (unspoken) challenge, then, is to provide an analysis of prevention that does not identify the prevented with what was going to happen but didn't. And, in the only critical discussion of Geach's view of which I am aware, Jonathan Kvanvig aims to do precisely that.[18] To Kvanvig's analysis I now turn.

(p.230) IV. Kvanvig on Prevention

As an analysis of prevention, Kvanvig offers us this:

> The truth of *p* is prevented by *S* =df. *S*'*s doing A* is causally sufficient in the circumstances for the falsity of *p*, and the circumstances apart from *S*'*s* doing *A* are such that *p* would have been true, were those circumstances to obtain.[19]

Kvanvig analyzes what it is for an *action* to prevent the truth of a *proposition*. However, we might wish for a more general account of prevention; while the sort of prevention at issue concerns human actions, presumably non-agential causes can still be preventative. Moreover, I find it more natural to think in terms of one *event* preventing another. Hence, I propose to replace Kvanvig's analysis with the following (compatible) analysis, where X and Y are *events*:

> X is prevented by Y = df. Y is causally sufficient for X's failing to obtain, and had Y not obtained, X would have obtained.

Let us call the proposed analysis *the counterfactual analysis of prevention*. The idea here is simple enough. On this analysis, the pilot's action prevented the crash because, had the pilot not acted, the plane would have crashed, and the pilot's action was sufficient for the plane's not crashing. Isn't this a satisfactory analysis of prevention? Geach identifies the prevented as that which was going to happen but did not. Kvanvig's counterfactual account identifies it as what would have happened had something else not.[20]

But the counterfactual account of prevention faces problems. First, it appears to lead to an explosion of preventions. Consider the following case. Presumably nearly everyone in Los Angeles went to bed at roughly the normal hour last night. Now, what would have happened had they not? Well, it is hard to say, but we can imagine that there would have been some additional car crashes today (on account of drowsy drivers) and perhaps some additional fights (on account of irritability), and a great deal else. But **(p.231)** everyone's going to bed at the normal hour was causally sufficient for these things not happening. Now, suppose some intrepid reporter at the *Los Angeles Times*, desperate for a story, and getting hold of the counterfactual account of prevention, pens the following:

> Actions of L.A. Residents Prevent Car Crashes and Fights
>
> In what would prove to be a fortunate turn of events, most L.A. residents went to bed at the normal hour last night, thereby preventing the great many car crashes and fights that would have resulted had they not.

This is not an altogether comfortable result for the counterfactual account of prevention. One might think it more appropriate to deny that L.A. residents going to bed at the normal hour prevented all this from happening. But the counterfactual account of prevention would have it that this is precisely what they did. Examples such as this could be multiplied *ad nauseam*, but the point is clear. A great many things that would have happened had something else not are not generally thought to have been *prevented* from happening. Geach has a diagnosis of why not: they were never going to happen in the first place.

Of course, this problem for the counterfactual account of prevention does not (so it seems to me) *decisively refute* the account. It is open to one to simply maintain that everyone's going to sleep at the normal time did prevent these things. But there are other potential problems for the view. Consider this example suggested to me by Kenneth Boyce:

> Nuclear war is not going to happen because our adversaries have firmly decided that they will not launch their nuclear weapons (in such a way as to causally settle the matter), regardless of whether Obama chooses to sign the peace treaty or not. But Obama does sign the peace treaty, and his doing so is also causally sufficient for nuclear war's not happening. However, in all the nearest worlds in which Obama does not sign the peace treaty, it is because McCain won the election. And in all those worlds, McCain launches *our* nuclear weapons, thereby causing nuclear war. So we have it that Obama does something that is causally sufficient for nuclear war's failing to occur, and we have it that had Obama not signed the treaty, nuclear war would have occurred. Yet it is false that Obama prevented nuclear war, since our adversaries had already decided not to launch the nuclear weapons regardless of whether he signs the treaty.

Geach, however, is not susceptible to this problem: given the firm intentions of our adversaries, nuclear war is not going to happen, and thus Obama does not prevent it by signing the treaty. But there are several features of this counter-example that deserve brief comment. First, it relies on the Lewis-Stalnaker account of the truth of counterfactuals, namely that **(p.232)** their truth is determined by so-called "nearness" to the actual world. Second, it exploits overdetermination: Obama's signing the treaty overdetermines that there is no nuclear war. Third, it relies on the truth of a *backtracker*: had Obama not signed the treaty, then McCain would have been elected (and would have launched our nuclear weapons). This last feature of the case is, of course, especially controversial; many would balk at the notion that this counterfactual is true. However, certainly many accept that such backtrackers are sometimes true, and thus this counter-example would seem to have some degree of purchase.

At any rate, what is clear is that the counterfactual account of prevention faces problems, problems the diagnosis of which seem to point towards Geachianism. And whether *any* counterfactual analysis of prevention can be successful is far from clear—especially so clearly successful as to make its adoption mandatory in the face of Geach's. However, if indeed no counterfactual account of prevention can be successful, then what would seem the most promising way of explaining away the appearance of a changing future is lost. Reconsider Geach's plane crash case. The plane was going to crash, but the pilot prevented this from happening. Geach sees here a changing future. But if one is averse to a changing future, by far the most natural thing to say about this case is simply that the plane would have crashed had the pilot not acted. However, it turns out that this counterfactual *does not* capture the fact that the pilot prevented the plane from crashing; prevention cannot be analyzed counterfactually. How then to explain away the appearance of a changing future? It isn't clear.

V. A Problem for Geach's Treatment of Prevention?

Suppose prevention cannot be analyzed counterfactually. That would certainly make it more difficult to explain away the force of Geach's examples. However, the mere failure (if it is a failure) of the counterfactual account does not by itself vindicate Geach's account of prevention. Geach's account may have problems of its own. Consider this case.[21] Suppose Jones is working on the uppermost levels of a skyscraper when he loses his balance and falls. Thankfully, he is caught by a safety net, which had been in place all along, and thus does not fall to his death. In this case, it is **(p.233)** natural to say that the safety net prevented his death. However, it does not seem appropriate in this case to say that Jones was going to fall to his death. He was never going to, precisely because the net was always there—the net was always going to prevent his fall. In this case, then, something prevented Jones from falling to his death, even though this was never going to happen. We thus seemingly have a counter-example to Geach's account of prevention.

What should the Geachian say in response to this example? I think the Geachian should argue that this case is not an instance of the phenomenon the analysis of which is currently at stake. The Geachian should say that what we are trying to analyze is what it is for one event to prevent another. But in this case, it is notable that it is the *net* that is said to have prevented Jones's fall to his death. And the net is not an event. Of course, one could attempt to translate this case into the parlance of events—perhaps the event of *the net's being there* prevented Jones's fall. But this is at best a gerrymandered event, and thus it is open to the Geachian to argue that the case is not an instance of the phenomenon in question.

But the Geachian can say more—much more. Suppose that workers on this particular building regularly fall and are caught by the net. Suppose, in fact, that they treat the net as just another feature of the building itself, which they perhaps jump on when they wish to quickly descend a few levels of the structure, and as it happens Jones merely slipped just prior to when he was about to jump in any case. It would then seem like the net prevents Jones's death in precisely the same manner in which the *floor* he is standing on prevents his death. For without the floor, of course, Jones would fall to his death. (Jumping on the net is, so the thought goes, just like stepping on the floor.) Of course, once we see prevention here, we'll see it *everywhere*. The couch I am sitting on prevents me from falling to the ground. The pillow supporting my laptop prevents it from falling onto my lap. And so on.

And if one claims that these cases are cases of the same sort of phenomenon Geach sought to analyze—that of one event preventing another—then the explosion problem not only returns, but returns with a vengeance. For it will turn out that (seemingly) an infinite number of events are now preventing my death. Which events? Well, for starters, *the oxygen in the room*'*s being there*, since without that oxygen, I would asphyxiate and die. Or suppose we have yet another intrepid reporter, this time at the *The New York Times*. And suppose he pens the following headline: "Events in New York today prevent the deaths of millions." What drama unfolded today in New York? Why, the braces of the skyscrapers in the city being there! Such a result is, of course, absurd. One could admit that the braces prevent the **(p.234)** collapse of the buildings, but deny that any events do so. But I do not see how this result can be avoided if one allows *the net*'*s being there* to be an event that prevented Jones's death. In sum, then, the Geachian may argue that an analysis of one event's preventing another is something worth having, and contend that cases of the sort in question pose no threat to Geach's analysis of this phenomenon.[22]

So let us recap. We began with a challenge to the Geachian: that all that was true, prior to the pilot's action, was *not* that the plane was going to crash, but that it was going to unless the pilot prevented it. Here, however, Geach complains that what is prevented *just is* what was going to happen, and thus that this way of explaining away the appearance of a changing future fails. Kvanvig objects that we can understand prevention counterfactually: we can understand the claim that the plane was going to crash but that the pilot prevented it (roughly) as the claim that it would have crashed had he not acted.[23] But I raised what I think are serious problems for the counterfactual account, problems the diagnosis of which point towards Geachianism. In sum, then, I do not think Kvanvig's objection decisively refutes Geach's view.

(p.235) VI. Kvanvig on Predictions

But Kvanvig has another criticism of Geach's view, which he nicely puts as follows:

> Before Johnny's operation, Geach correctly notes that it was right to say "Johnny is going to bleed to death." However, what is not clear is that the statement *Johnny is going to bleed to death* is true before the operation. From the perspective of the observers, it appeared true and hence it was proper to claim to be true; but something can be proper to claim even though untrue. In this case we can appreciate the difference by noting that it would be appropriate for the observers to note, after the operation, that their earlier claim was false[24]

By way of response, we first ought to note that Geach's view need not have it that *every time* something appears to us as if it is going to happen, it is true that it *is* going to happen. Thus, it could very well be that (despite appearances) Johnny was not going to bleed to death all along, and thus it would be appropriate for the observers to note after the operation that their earlier claim was false. Despite what Geach may intimate, his examples are best thought of as what is *in principle possible*, and not as examples of how things always are.

But suppose this is a case in which it is indeed taken to be true that Johnny was going to bleed to death.[25] Then the response to Kvanvig must be hard-line. In this case, it is not appropriate for the observers to note after the operation that their earlier claim was false. Rather, it is appropriate for them to be thankful that the future changed. Suppose Rob unexpectedly loses his job. Should he conclude that he was going to lose it all along—that it was true two weeks ago, when he was working on his crucial project, that he was going to lose it? Of course not, says the Geachian. Perhaps he *was* **(p.236)** going to keep it, but disaster struck his project. In the present case, should we conclude from the fact that Johnny has not bled to death that he never was going to? Again, of course not.[26]

It is well known that views denying that future contingents are true face the problem of accounting for the practice of retroactively predicating truth to predictions. This is Kvanvig's criticism; since Johnny has not in fact bled to death, it follows that it would have been true to say before the operation that he would not, and consequently that it was false that he would.[27] It must be admitted, I think, that this practice does constitute some evidence against such views. Geach's view faces the same problem. Various strategies of explaining away such retroactive predications of truth have been proposed, and here I simply note that such (or similar) strategies are likewise available to the Geachian.[28] In short, I do not believe this problem to be decisive against the Geachian. Thus, while Kvanvig's criticism of Geach's proposal does point out a legitimate problem for his approach, Geachianism still (I claim) emerges as a theoretical contender.

VII. Smart and Plantinga on the Impossibility of Changing the Future

Besides the logic of prevention, there is at least one other argumentative strategy one may use to reach Geach's conclusions, a strategy recently pursued by **(p.237)** Mark Hinchliff.[29] According to Hinchliff, it is commonplace and intuitive that we cannot change the past, but that we can change the future. It is a virtue of a theory if it can preserve this asymmetry. But (arguably) only on Geach's view can we genuinely change the future.

But it is here that we encounter incredulous dismissals of a changing future by J. J. C. Smart and Alvin Plantinga.[30] Here is Smart:

> It makes no more sense to talk of changing the future than it does of changing the past. Suppose that I decide to change the future, by having coffee for breakfast tomorrow instead of my usual tea. Have I changed the future? No. For coffee for breakfast *was* the future…. [T]he fact that our present actions determine the future would be most misleadingly expressed or described by saying that we can change the future. A man can change his trousers, his club, or his job…. But one thing he cannot change is the future, since whatever he brings about *is* the future, and nothing else is, or ever was.[31]

Not surprisingly, Geach regards Smart's statements here as mere assertion. After giving us Smart's quote, Geach immediately says:

> “Or ever was,” indeed! If A is bringing something about, no doubt that is what *is now* going to happen; but what has Smart done to show that it *always was* going to happen, even before A's action? Nothing; he has merely asserted it.[32]

Plantinga says even less than Smart. Can someone, say Paul, change the future? Says Plantinga:

> To alter the future, Paul must do something like this: he must perform some action *A* at a time *t* before 9:21 such that prior to *t* it is true that Paul will walk out at 9:21, but after *t* (after he performs *A*) false that he will.

This is, of course, precisely what Geach believes can happen. But Plantinga says, “neither Paul nor anyone—not even God—can do something like that. So the future is no more alterable than the past.”[33] The Geachian diagnosis is again clear: this is mere assertion.

(p.238) In Smart's case, the fundamental disagreement clearly arises from the fact that Smart is a *reductionist* about tense or a so-called *de-tenser*.[34] Such a view is plainly at odds with Geach's. According to the reductionist, all tensed statements are to be analyzed in purely *tenseless* terms; tense is not a fundamental feature of reality. On such a view, propositions have their truth values *eternally*, so it is impossible for a proposition to have one truth value at one time and a different one at another; in fact, propositions are not true at times at all, but simply *true*. Such a view is closely related to the so-called B-theory of time. According to the B-theory, there is no objective present or “now.” Events are not objectively past, present, or future, but are merely earlier than, simultaneous with, and later than one's vantage point. And events have their B-series order *eternally*—if X is ever earlier than Y, then X is *always* earlier than Y.

Geach, however, is working with a rival conception of time and tense.[35] Unlike Smart's, Geach's is a *tensed* theory of time—he “takes tense seriously,” as it is sometimes said. On this view, tensed features of our statements are not reducible to, and cannot be understood in terms of, tenseless statements. Propositions are true *at times* and can change truth values over time. The tensed view is closely related to the A-theory of time. On the A-theory, there *is* an objective present, and events have irreducibly tensed properties of being past, present, or future. To be sure, a tensed theory of time does not straightforwardly *entail* Geach's position. But it is important to see whether objections to Geach's view are in fact objections to the tensed theory of time more generally, which has many able defenders, including, of course, Geach himself.[36]

So much for Smart and Plantinga on changing the future. But one last point. On Geach's view, even though something *will* happen—*actually* will happen—it can very well fail to happen. Perhaps you think this is a strange, counterintuitive result. If it really is true that something is going to happen, how could it fail to? If this has been your reaction to Geach's view, then **(p.239)** I must warn you that you may be in danger of becoming an open futurist. For it is clearly the central thesis of open-future views—and logical fatalism—that if something will happen in the future, it *cannot* later fail to happen. If it will happen, then it is inevitable. Some of Geach's in-house rivals are welcome to feel incredulity at this feature of Geach's view. Those who do not wish to enter the fold are not.[37]

This is, I believe, an important issue. Indeed, commenting on an earlier draft of this paper, one friend (and a very good philosopher!), himself certainly no friend to open-future views, fell hard for the "trap" I had wanted to set up above. That is, he expressed wholehearted agreement that the relevant claim—that what will happen can fail to happen—is a "strange, counterintuitive result." But moments later he noticed where that had led him: away from Geach, but now into the arms—or the jaws—of the fatalist. Of course, this lone anecdote hardly *establishes* anything. But I believe it reveals something important. Geach presents us with a view on which what will happen might not. Open futurists (and fatalists) have an easy time criticizing this view. Others, however, must tread carefully. They must claim that what is objectionable about Geach's view is his claim that what will happen might—well, what? Not *might not*. Perhaps *might later be such that it won't*. But I think it is implausible to suppose that intuition discriminates so finely between these claims; in the case of my friend, it clearly did not—at least initially. That is, I claim that if you find the idea of a changing future counterintuitive, it may very well be because (deep down, in the fatalist recesses within us all) you find the claim that what will happen nevertheless might not be counterintuitive as well.[38] If so, then we have some ironic results: in expressing what is wrong with one version of open theism, one expresses what is right about another. No doubt many would resist this claim, and maintain that we can perfectly well see that Geach's view is wrong, but not because we do not like his result that what **(p.240)** will happen might not happen. I'm skeptical. But I here merely register my skepticism.[39]

This must end our discussion of Geach's view of the mutability of the future, which, needless to say, raises many fascinating questions that we cannot address here.[40] A new issue now arises. How does Geach's view bear on the compatibility of divine foreknowledge and human freedom? And how does the application of the view to this issue relate to other versions of open theism? To answer these questions, we first need to investigate the argument to which open theism has largely been a reaction. In the minds of most, especially those coming at this topic from a philosophical angle, open theism is primarily a particular sort of response to the argument for the incompatibility of divine foreknowledge and human freedom, an argument that was crystallized by Nelson Pike in his 1965 paper, "Divine Omniscience and Voluntary Action." I thus turn to a brief discussion of this argument, after which we will see how Geach's view opens up an interesting new response to it.

VIII. Geach and the (in?)Compatibility of Foreknowledge and Freedom

Pike's 1965 paper provoked a veritable avalanche of work on the problem of divine foreknowledge and human freedom. According to Pike, one can argue from certain plausible premises together with the thesis that God exists to the conclusion that no human action is free (in the sense requiring the ability to do otherwise). Here, it is worth simply replicating Pike's argument in full:

1. “God existed at $t1$” entails “If Jones did X at $t2$, God believed at $t1$ that Jones would do X at $t2$.”
2. “God believes X” entails “X is true.”
(p.241)
3. It is not within one's power at a given time to do something having a description that is logically contradictory.
4. It is not within one's power at a given time to do something that would bring it about that someone who held a certain belief at a time prior to the time in question did not hold that belief at the time prior to the time in question.
5. It is not within one's power at a given time to do something that would bring it about that a person who existed at an earlier time did not exist at that earlier time.
6. If God existed at $t1$ and if God believed at $t1$ that Jones would do X at $t2$, then if it was within Jones's power at $t2$ to refrain from doing X, then (1) it was within Jones's power at $t2$ to do something that would have brought it about that God held a false belief at $t1$, or (2) it was within Jones's power at $t2$ to do something which would have brought it about that God did not hold the belief He held at $t1$, or (3) it was within Jones's power at $t2$ to do something that would have brought it about that any person who believed at $t1$ that Jones would do X at $t2$ (one of whom was, by hypothesis, God) held a false belief and thus was not God—that is, that God (who by hypothesis existed at $t1$) did not exist at $t1$.
7. Alternative 1 in the consequent of item 6 is false. (from 2 and 3)
8. Alternative 2 in the consequent of item 6 is false. (from 4)
9. Alternative 3 in the consequent of item 6 is false. (from 5)
10. Therefore, if God existed at $t1$ and if God believed at $t1$ that Jones would do X at $t2$, then it was not within Jones's power at $t2$ to refrain from doing X. (from 6 through 9)
11. Therefore, if God existed at $t1$, and if Jones did X at $t2$, it was not within Jones's power at $t2$ to refrain from doing X. (from 1 and 10)[41]

It is beyond the scope of this paper to detail all the objections philosophers have had with Pike's argument and all the sorts of replies it has provoked. Notably, however, perhaps the most common reply has been the so-called “Ockhamist solution” advocated by (amongst others) Marilyn Adams and Alvin Plantinga.[42] The Ockhamist denies premise (4) of Pike's argument on grounds that God's beliefs are “temporally relational” or “soft” facts, and thus not subject to the (tacit) principle of the fixity of the past underlying **(p.242)** (4).[43] That is, by the Ockhamist's lights, we can sometimes so act that God would not have believed what he actually did believe.

The open theist does not take this route. The open theist's response to Pike's argument is to deny (1), the thesis that it follows from Jones's doing X at *t*2 that God believed (and hence knew) at *t*1 that Jones would do X at *t*2. Both versions of open theism discussed in Rhoda's analysis make this reply salient. On one version, it was not *true* at *t*1 that Jones would freely do X at *t*2, and on the other, it was logically impossible (for reasons Pike's argument—setting aside certain niceties—itself brings out) for God to believe at *t*1 that Jones would freely do X at *t*2. But now we can see how Geach's view also makes possible a denial of (1). Recall Geach's rhetorical question:

> But if something did happen, doesn't this show it was after all going to happen? Certainly; but not that it *always* was going to happen.

In other words, Geach denies the inference from X's happening at *t*2 to its being true at all earlier times that X would happen at *t*2. Thus, it does not *follow* from Jones's doing X at *t*2 that it was true at *t*1 that Jones would do X at *t*2. At *t* 1, it may have been true that Jones was going to do ~X at *t*2 rather than X, and in the time between *t*1 and *t*2 prevented what was going to happen (his refraining from X-ing at *t*2) and did X at *t*2 instead. Moreover, the Geachian may plainly maintain that if it was not true at *t*1 that Jones would do X at *t*2, then God did not believe at *t*1 that he would do X at *t*2, contrary to Pike's first premise.[44]

On Geach's analysis, it could very well be that at *t*1 God believes of Jones that he will do X at *t*10, but at *t*7 something happens (perhaps a decision of Jones himself) so as to prevent Jones from doing X at *t*10, with the result that Jones does not do X at *t*10, contrary to what at *t*1 God believed he would do. Someone in the grips of a view opposed to Geach's would immediately infer from this result that God was *mistaken* at *t*1 in his belief about Jones doing X at *t*10. But such an inference would be the real mistake, according to the Geachian. Rather, if God believes at *t*1 that Jones is going to do X at *t*10, then at *t*1 it of course really is the case that Jones is going to do X at *t*10; God is making no mistake. Nevertheless, it is consistent with its being true at *t*1 that Jones will do X at *t*10 that at *t*7 it becomes false that **(p.243)** Jones will do X at *t*10. In this case, while God begins believing at *t*7 that Jones will not do X at *t*10, he has not always believed this; previously, he believed (correctly) that Jones would do X at *t*10, but adjusted his beliefs about what would happen to accommodate the fact that what was going to happen has been prevented from happening, and a new thing now will happen. Whether this picture of God's beliefs is plausible is an open question, but there is clearly no basis for a charge that God would be making mistakes.

But Geach's view opens up another way of reply to Pike's argument that is somewhat more complicated than a simple denial of (1). Geach may deny that the trichotomy in (6) is exhaustive.[45] On Geach's view, Jones's power to refrain from doing X at *t*2 (despite God's belief at *t*1 that he would do X at *t*2) need not be the power so to act that God either would have held a different belief at *t*1, a false belief at *t*1, or would not have existed at *t*1. Rather, Jones's power could simply be the power so to act that though God believed at *t*1 that Jones would do X at *t*2, what God believed would happen fails to happen. In other words, by preventing what was going to be (as of *t*1) from coming to pass, Jones brings it about that what God believed would happen was prevented from happening. Again, it is tempting here to insist that this power of Jones's is the power to bring it about that God was mistaken, but the Geachian denies this. Moreover, Jones's power is *not* the counterfactual control over the past invoked by Ockhamists that open theists have traditionally found problematic. Jones *does not* have the power so to act that God's beliefs would have been different, as the Ockhamist claims. Rather, he has the power so to act that God's (present) beliefs *become* different.

IX. Geachianism and Open Theism

As we just saw, given Geach's view, one may maintain that Jones has the power to refrain from doing X at *t*2 (despite God's belief at *t*1) without invoking any problematic power over the past. Given this result, it is worth briefly pausing to consider how this result squares with open theism commonly conceived. Suppose we are at *t*20, a time later than *t*10. Using the past-tense construction to avoid certain difficulties about a tenseless “happens” or “does,” notice that Geach's view maintains that:

(A) God believed at *t*1 that Jones would do X at *t*10

(p.244) does not entail that:

(B) Jones did X at *t*10.[46]

Now, if it does not follow from God's belief at *t*1 that Jones would do X at *t*10 that (looking backwards) Jones *did* X at *t*10, it will be hard to see how it could follow from God's belief that Jones would do X at *t*10 that Jones *could not have done otherwise* than X at *t*10. Given Geach's views about future contingents, it will be impossible to generate an argument for the incompatibility of mere divine foreknowledge and human freedom.[47] Even if God knew that Jones would do X at *t*10, Jones could have (*ceteris paribus*) acted so as to prevent its being the case that he did X at *t*10—again, even if he did X at *t*10 and God knew he would.

Given this result, we can notice that Geachianism is in an important respect significantly different than others versions of open theism. After all, while all open theists deny (1) of Pike's argument, the claim that it follows from Jones's doing X at *t*2 that God believed at *t*1 that Jones would do X at *t*2, open theists are typically taken to maintain that if God *really did* believe of Jones that he would do X at *t*2, then of course it would follow that Jones cannot refrain from doing X at *t*2. In other words, the traditional motivation for open theism has been the thought that divine foreknowledge and human freedom are incompatible, with the consequence that we must deny that God has foreknowledge of future free actions. Though Geach's view entails a denial of (1), it also has the consequence that mere divine foreknowledge is consistent with human freedom.

Some might take this fact to call into question the claim that Geachianism is a version of open theism. Of course, I have no decisive argument against the view that the incompatibility thesis is essential to open theism. And anyway, so long as we are clear about how we use our terms, what we call the relevant views is not of ultimate concern. However, it strikes me that Geachianism *is* a version of the view, and that what makes it so is its denial **(p.245)** of premise (1) of Pike's argument: from Jones's doing X at *t2*, it follows that at *t*1 God believed he would. I submit that open theists must deny this claim. More generally, I submit that you are an open theist if and only if you maintain Rhoda's first two criteria (theism and indeterminism), together with the following:

> (OT) Some things happen and have happened which God has not always known would happen.

Notice that this thesis is directed towards the *present* and the *past*, and not the future. This makes sense, since the disagreements that mark different versions of open theism are disagreements about the nature of the *future*.

So let us come full circle and reconsider Rhoda's characterization of “generic open theism.” Recall that Rhoda maintained that open theists must endorse the claim that there are some states of affairs such that God knows neither that they *will* nor *will not* happen. But we can see that the Geachian need not admit this. It is not essential to Geachianism that there are some things about which God must say, “I do not believe that that is going to happen, and nor do I believe that it is not.” Admittedly, this would imply that God has in mind some unique, maximal blueprint of how the world will be. But how the world will be changes. The result is a denial of premise (1) of Pike's argument, and thus, I suggest, a version of open theism. So I suggest replacing Rhoda's analysis of “generic open theism” with the one suggested above.[48]

X. Geachianism and Divine Repentance Texts

In the final two sections of this paper, I consider how Geachianism bears on two distinctively theological considerations historically relevant to open theism: so-called "divine repentance texts" and the doctrine of providence. Consider this passage from Isaiah, which, along with others like it, is often cited by open-theist theologians: **(p.246)**

> In those days Hezekiah became ill and was at the point of death. The prophet Isaiah son of Amoz went to him and said, "This is what the LORD says: Put your house in order, because you are going to die; you will not recover."
>
> Hezekiah turned his face to the wall and prayed to the LORD, "Remember, O LORD, how I have walked before you faithfully and with wholehearted devotion and have done what is good in your eyes." And Hezekiah wept bitterly.
>
> Then the word of the LORD came to Isaiah: "Go and tell Hezekiah, 'This is what the LORD, the God of your father David, says: I have heard your prayer and seen your tears; I will add fifteen years to your life.' " (38:1–5, New International Version)

According to open theologians, a straightforward reading of this and similar passages puts pressure on the traditional model of divine foreknowledge. After all, so the thought goes, if God eternally and infallibly knew that Hezekiah was not going to die, it would seem that God was being disingenuous when he told Hezekiah that he would.

Perhaps open theologians are right in seeing a problem for the traditional model here. What is remarkable, however, is that a straightforward reading of the passage—and that is what the theologians in question claim to be interested in—supports *Geachianism* rather than familiar models of open theism. For only on Geachianism can God's statements to Hezekiah come out *true.* Hezekiah really *was* going to die, and God was *correct* in saying that he would not recover. But Hezekiah's earnest prayer *prevented* what was going to happen, so that what was once true (that he would die) now becomes false. Indeed, a Geachian of a certain sort might say that just as such texts pose a stark problem for the traditional model (on account of putting a known falsehood in God's mouth), they pose the very same problem for traditional open views. Traditional openness theologians must explain how God could have said that Hezekiah would die, even though this was (by their lights) not true. At any rate, a Geachian approach to such passages seems required if we really are to take them at face value and preserve the notion that God is no deceiver. I think it is remarkable that this fact has not been widely noticed. Perhaps I have just provided theologians of a certain inclination good reason to become Geachians.

Moreover, even if there is no good reason to want (let alone *require*) “straightforward” readings of such texts (or even if one accords no even *prima facie* authority to such texts in the first place), the text in question still seems to help the Geachian. For what would be so problematic about reading the text in the straightforward way? It is not clearly *incoherent* or in any obvious sense *impossible* that the text could be representing things as they really are. Indeed, the text seems to be a perfect illustration of Geachianism in action. Of course, those in the grip of the traditional view would have to **(p.247)** insist that God was merely teaching Hezekiah a lesson (or some such), and that he knew all along that his (God's) message to Hezekiah would lead to Hezekiah's earnest prayer and his “adding” fifteen years to Hezekiah's life. But why is an explanation along these lines *required*? Why not rather say that Hezekiah (or God) changed the future?

XI. Geachianism and Providence

Perhaps the easiest way to consider how Geachianism might affect one's theory of providence is to consider one at least *prima facie* difficulty a Geachian model of providence faces. It may seem as if Geachianism would imply that God cannot count on hardly *anything's* happening in the future, and if so, this fact would seem to call into question God's ability to providentially govern the world as required. After all, Geachianism has the result that God could know that X will happen, but that it (causally) could still fail to happen. Moreover, God would not have access to some immutable fact about whether X “happens” at the relevant time. Knowing that something will happen by itself provides no guarantees that it later will not.

Is this a severe problem for Geachianism? Not obviously. Geachianism would seem to call for a distinction between those things that *must* happen and those things that are merely *going to* happen. If God were wise, he would know to build his plans—at any rate, the really important ones—only on the foundation of those things that *must* happen. Whereas open-future open theists identify what *will* happen with what *must* happen, Geach would deny this. Nonetheless, a Geachian could agree that there are some things that are now determined, and could maintain that these are those things that must happen. Just as on open-future open views God would be unwise to base some crucial plan on what merely will *probably* be, so the Geachian could insist that God would be unwise to base some crucial plan on merely what will be. For that could always change. At any rate, what is clear is that Geach's God will be in the exact same providential situation as traditional open views suggest. So if traditional open views have an adequate account of providence, so does Geachianism.

Moreover, it is at this point that I believe Geachianism is able to shine considerable light on controversial issues concerning the providential usefulness (or lack thereof) of so-called “simple foreknowledge.” As I remarked above, the Geachian God could know that X will happen, but it also be true that X could (causally) still fail to happen. Of course, Ockhamists too must **(p.248)** accept this thesis, but the distinctions here are obvious: the Ockhamist God would still know that X *does unchangeably* happen, even if it causally might not. This knowledge might seem to give the Ockhamist God a providential advantage over the Geachian God (or any God lacking such knowledge). But William Hasker has vigorously—and, in my judgment, correctly—argued that this knowledge is in fact providentially *useless* to God.[49] For on the Ockhamist view, facts about what unchangeably happens at times are facts logically and explanatorily *posterior* to God's deliberations (or other providential activities). Thus, God cannot *use* knowledge of what (unchangeably) happens as data in deciding what the future is to be like; by that (logical) time, the future is already unchangeably fixed.

Suppose, as I believe to be the case, that Hasker is right in holding that (given Ockhamist suppositions) knowledge of what will be is providentially useless to God. This would seem to be a cost for Ockhamism. For it is counterintuitive to suppose that knowledge of what will be would be useless in controlling the world as one sees fit. Plausibly, however, only Geachianism can properly accommodate this intuition. That is, on Geachianism, knowledge of what will be is *clearly* providentially advantageous: by knowing what will happen, one is thereby in much better position to prevent it should one wish to do so. Or, if something one *does* want to happen (at the moment) *will not* happen, then one can then act (or try to act) so as to bring it about that it will. Common sense has it that total knowledge of what is going to happen would be a great help to someone wishing to control the future. Traditional models of foreknowledge cannot accommodate this result. Geachianism can. Though these points clearly deserve further elaboration, I believe they constitute an independent argument for Geachianism.[50]

(p.249) XII. Conclusion: Advantage to the Geachian?

Here ends my application of Geach's view to the issues of divine foreknowledge and providence. The view that emerges is unique, and it is worth pausing to consider whether it has distinctive advantages over its more familiar rivals. It does. First, consider open-future versions of open theism. One such version denies bivalence. This has been seen to be a severe cost for the view. Geachianism requires no rejection of bivalence. The other such version maintains that *will* and *will not* are not contradictories, so that “Jones will swim tomorrow, or Jones will not swim tomorrow” is not an instance of $p \vee \sim p$. This has struck many as dubious, and at any rate is not the view of the overwhelming majority. Geachianism does not deny that *will* and *will not* are contradictories. Next, consider Hasker's limited-foreknowledge version of open theism, the version maintaining that there are truths about the future that God does not know. Hasker and company must explain how it is that God can still be omniscient, despite not knowing these truths. The Geachian faces no such problem. These are the most familiar problems for these versions of open theism. Geachianism avoids them all.

Moreover, looking beyond its advantages over other open views, many have thought the Ockhamist's counterfactual control over the past to be problematic. Not a problem for the Geachian. Nor does the Geachian have any problems arising from God's supposed atemporality. There are no counterfactuals of freedom in need of grounds. One has a straightforward account of the providential usefulness of knowing what will happen. All this can be bought by maintaining that the future is mutable. But what is the cost? As Kvanvig's critique showed, one has the problem of accounting for retroactive predications of truth to predictions. But this cost does not strike me as decisive against open views of the future. Perhaps there are other problems for the Geachian, but they must fall outside the scope of this paper.

In this paper, I have tried to develop Geach's view on future contingents and apply it to the problems of divine foreknowledge, providence, and human freedom. Geach's work is provocative, but he did not systematically explain how his views are related to and are different from rival views about time, tense, foreknowledge, and freedom. I have sought to do so here, albeit in a preliminary way. I do not claim to have considered—let alone explored—all the various problems that might face the Geachian. However, I do claim that philosophers should be aware of Geach's view as a theoretical possibility in these debates. Geach's case from the logic of prevention is stronger than the arguments of *Providence and Evil* might make it seem. Kvanvig's criticisms are not decisive. The quick dismissals of Smart and **(p.250)** Plantinga are too quick. And Geachianism has certain attractive advantages over its rivals. The dialectical terrain about such matters very rarely sees a genuinely new position emerge. Perhaps the dialectic was going to proceed as it has been for some time to come. If so, then in a bold Geachian spirit, I hope to have changed the future.[51]

References

Bibliography references:

Belnap, Nuel and Green, Mitchell (1994), “Indeterminism and the Thin Red Line,” *Philosophical Perspectives*, vol. 8, in *Logic and Language*, James Tomberlin (ed.) (Atascadero, CA: Ridgeview Publishing Co.), 365–88.

Dowe, P. (2006), “A Counterfactual Theory of Prevention and ‘Causation’ by Omission,” *Australasian Journal of Philosophy* 79: 216–26.

Dummett, Michael (1982), “Realism,” *Synthese* 52: 1.

Fischer, John Martin (ed.) (1989), *God, Foreknowledge, and Freedom* (Stanford, CA: Stanford University Press).

—— Todd, Patrick, and Tognazzini, Neal (2009), “Re-reading Pike's ‘Divine Omniscience and Voluntary Action,’ ” *Philosophical Papers* 38: 2 (July 2009): 247–70.

Fredosso, Alfred (1983), “Accidental Necessity and Logical Determinism,” *Journal of Philosophy* 80: 257–78, reprinted in Fischer (1989: 136–58).

—— (1988), “Introduction” to Luis de Molina, *On Divine Foreknowledge: Part IV of the Concordia*, tr. Alfred J. Fredosso (Ithaca, NY: Cornell University Press).

Geach, Peter (1965), “Some Problems about Time,” *Proceedings of the British Academy* 11.

—— (1973), “The Future,” *New Blackfriars* 54, 208–18.

—— (1977), *Providence and Evil* (Cambridge: Cambridge University Press).

—— (1998), *Truth and Hope* (Notre Dame, IN: University of Notre Dame Press).

Hartshorne, Charles (1965), “The Meaning of ‘Is Going to Be,’ ” *Mind* 74: 293: 46–58.

Hasker, William (1989), *God, Time, and Knowledge* (Ithaca, NY: Cornell University Press).

(p.251) Kenny, Anthony (1987), *The God of the Philosophers* (Oxford: Oxford University Press).

Kvanvig, Jonathan (1986),*The Possibility of an All-knowing God* (New York: Macmillan).

Lucas, John Randolph (1989), *The Future: An Essay on God, Temporality, and Truth* (Oxford: Basil Blackwell, 1989).

Pike, Nelson (1965), “Divine Omniscience and Voluntary Action,” *The Philosophical Review* 74: 27–46, reprinted in Fischer (1989: 57–73).

Plantinga, Alvin (1986), “On Ockham's Way Out,” *Faith and Philosophy* 3: 235–69, reprinted in Fischer (1989: 178–215).

Prior, A. N. (2003), “The Formalities of Omniscience,” *Papers on Time and Tense* (Oxford: Oxford University Press).

Rhoda, Alan (2008), “Generic Open Theism and Some Varieties Thereof,” *Religious Studies* 44: 225–34.

—— Boyd, Gregory, and Belt, Thomas (2006), “Open Theism, Omniscience, and the Nature of the Future,” *Faith and Philosophy* 23: 432–59.

Smart, J. J. C. (1964), “Introduction,” in his *Problems of Space and Time* (New York: Macmillan).

Swinburne, Richard (1993), *The Coherence of Theism* (Oxford: Oxford University Press).

Van Inwagen, Peter (2008), “What Does an Omniscient Being Know About the Future?” *Oxford Studies in Philosophy of Religion* vol. 1, Jonathan Kvanvig (ed.) (Oxford: Oxford University Press). **(p.252)**

Notes:

(1) This is a deficiency my co-authors (John Martin Fischer and Neal Tognazzini) and I have recently tried to remedy. See Fischer et al. (2009). We give a brief sketch of the Geachian view that I here develop at length.

(2) However, just as there is sometimes a question as to whether Aristotle is an Aristotelian, there may likewise be a question whether Geach is a Geachian. As I remarked, Geach did not systematically develop his position on these matters. Thus, while the position I develop seems to me to be the plain consequence of what Geach suggests, in what follows I may go beyond what Geach himself would endorse.

(3) See Rhoda (2008).

(4) See Kvanvig (1986: ch. 1).

(5) The relationship between process theism and open theism is historically complex, but I need not enter into the relevant details. In short, open theists have consistently wished to distinguish their views from those of process theists, who *inter alia* deny creation *ex nihilo* and God's ability to unilaterally intervene in the world in the ways imagined by traditional theists.

(6) See, e.g., Rhoda et al. (2006).

(7) Though Aristotle is widely interpreted as denying bivalence for future contingents in *On Interpretation 9*, some philosophers—notably Anscombe—have disputed this.

(8) See Lucas (1989).

(9) By including Hartshorne here, I do not mean to imply that Hartshorne is himself an open theist; he is, rather, a process theist whose views on future contingents are amenable both to process theism and open theism. Hartshorne was the first (of which I am aware) to defend the view that future contingents are uniformly false. According to Hartshorne (and later Rhoda), if at *t*1 it is not determined that X happens at *t*3 (or does not happen at *t*3), then at *t*1 it is false that X will happen at *t*3 and it is false that X will not happen at *t*3. “Will” and “will not” are *contraries* rather than *contradictories.* Even if the view in question entered contemporary discussion primarily through Prior (Rhoda's primary source for the view), as seems to be the case, it would be unseemly not to credit Hartshorne here. Even if he is not an open theist, his view can be (and has been) adopted by open theists. See Hartshorne (1965) for his development of the view.

(10) See Rhoda et al. (2006). See also Prior (2003). Prior seems to *formulate* the view in question (see especially p. 54's discussion of Peirce), but does not unambiguously endorse it, and in places would seem to deny bivalence. Prior's work on this matter is complicated, and for the purposes of this paper I set aside the question of his actual views.

(11) See Swinburne (1993); Hasker (1989); and Van Inwagen (2008).

(12) Belnap and Green (1994).

(13) Some may dispute this claim. In its defense, I merely note that it is plausible to suppose that if something is going to happen, then it will happen, and conversely, that if something will happen, it is going to. I return to this issue in fn. 26.

(14) Geach (1977: 47).

(15) Ibid. (48).

(16) Ibid. (50).

(17) Geach (1977: 51–2).

(18) There is a brief statement of Geach's view in Anthony Kenny (1987: 53–4). There is also a discussion of Geach's view of the future in the Introduction to Fischer (1989: 23–5). In short, Fischer—taking after Kvanvig—represents Geach as claiming that truths about the future are only *apparently* about the future, and are really about the present tendencies of things. But these issues are orthogonal to my purposes in this paper; I thus set them aside. Also, Alfred Fredosso (1986) has a brief mention of the view, as does Michael Dummett (1982: 87).

(19) Kvanvig (1986: 10).

(20) Other philosophers have sought to analyze prevention counterfactually. The literature on this topic has as its primary origin attempts to understand the nature of *causation*. That is, the central questions are, e.g., whether preventions can be causes, whether by catching a ball that was headed towards a window, I prevent the window's being broken, even if you would have caught the ball anyway, and so on. See, e.g. Dowe (2001). I believe Dowe's analysis, while slightly different than Kvanvig's, is subject to the same problems.

(21) Thanks to an anonymous referee for pressing this case.

(22) More generally, of course, the Geachian should (I say) contend that in those cases in which it *is* correct to say that one (non-gerrymandered) event prevented another, it will also be the case that the prevented event was going to happen. However, I hesitate to tie the success of Geachianism too closely to Geach's claim that what is prevented is always what was going to happen. What would it mean for Geachianism if this case (or some other) were a counter-example to this claim? Recall the initial challenge: that all that was true prior to the pilot's action was that the plane was going to crash unless prevented from crashing. Again, here Geach says: but what is prevented just is what was going to happen. But this reply is not open to Geach if there are counter-examples to this thesis. And if so, then Geachianism certainly loses some strength; part of its strength comes from the availability of this Geachian reply and its forcing the objector to take up the seriously problematic counterfactual account of prevention. So I would recommend to any Geachian a stalwart defense of Geach's account of prevention. However, a Geachian *could* simply insist that even if one could maintain that only the mere conditional was true, there is no ultimately compelling reason to adopt this thesis. That is, perhaps someone will say that all that was true was that the plane was going to crash unless prevented from crashing. However, the Geachian may simply have the firm intuition that this *isn't* all that was true: it was also true that the plane was going to crash. If so, he needn't rely on the reply that what is prevented just is what was going to happen (though, again, this claim would certainly help his case). And thus he needn't defend this claim against (putative) counter-examples. In short, he simply needs the claim that prevention *sometimes* changes the future, not that it always does.

(23) Note: the Geachian's claim is not (or needn't be) the claim that this counterfactual is not *true*, but that it does not account for the case being one of prevention.

(24) Kvanvig (1986: 7).

(25) These points raise questions for the Geachian. Just when is it true—and what, if anything, makes it true—that something is “going to happen”? This is an extremely difficult question, and I do not have the space to adequately address it. I note only that all those philosophers who have commented on Geach's view have attributed to him— and not for no good textual reason—the view that X is going to happen iff there is some appropriate present tendency towards X's happening. While such a view is an option for the Geachian, it is not the only one: the Geachian can maintain that what is going to happen is simply brute—that what, if anything, makes it true that X is going to happen is the fact that it is going to. Ironically, it seems plain to me that the only theorists in this debate who can mount a principled objection to this claim are other open theists—open futurists who claim that such truths must be grounded in determinative causal tendencies. These points deserve further elaboration, like others in this paper, but I must set them aside.

(26) Perhaps some will complain that I am cheating here by framing the issue in terms of what was *going to* happen rather than in terms of what *would* happen. For some may say that it is more counterintuitive to suppose that something *would* happen but later does not than it is to suppose that something *was going to* happen but later does not. Well, perhaps I am indeed cheating here. However, at this stage of the dialectic, I believe the Geachian is entitled to the equivalence of *X is going to happen* with *X will happen*. Perhaps these expressions mean different things or have different truth conditions. If so, then perhaps this raises a problem for the Geachian, or makes the present problem more difficult. But I do not see that it is dialectically infelicitous to rely on this equivalence in the current defense of the view.

(27) For another (later) statement of this objection, see Fredosso (1988: 71–2).

(28) See, e.g. Rhoda (2006). Note: I am not claiming that the problem Kvanvig has raised for Geach is entirely isomorphic to the similar problem facing open future open theists. For on Geach's view, from the fact that X has not happened, we cannot conclude that it was never going to—“X will happen” might have been true at some time in the past. Not so for open futurists. Still, *both* maintain that from the fact that X *has* happened, we cannot conclude that it was always going to. However, it is not clear that this difference makes the prediction problem any harder for the Geachian than it is for the open futurist; at any rate, I set this issue aside.

(29) As far as I know, Hinchliff is the only contemporary philosopher who defends Geach's position. Hinchliff defended the view in a talk given at the University of California, Riverside in the fall of 2007. Indeed, I owe it to Hinchliff for bringing Geach's view to my attention. A defense of the Geachian position has not yet appeared in Hinchliff's published work.

(30) To be clear, neither Smart nor Plantinga are explicitly addressing *Geach* in these passages.

(31) Geach (1977: 52).

(32) Ibid. (52).

(33) Plantinga (1986), reprinted in Fischer (1989: 189). I thank Mark Hinchliff for bringing this passage to my attention.

(34) See Smart (1964).

(35) See, e.g., Geach's important and influential 1965 essay, “Some Problems about Time.”

(36) There are important related issues here regarding the *ontology* of time that deserve mention. Eternalism is the view that past, present, and future objects all equally exist. Geachianism would seem plainly to require non-eternalism; surely, that is, we are not to imagine that when the pilot prevents the crash, he brings it about that the concrete mangled ruins of a (future) plane pop out of existence. As Geach says (somewhat cryptically), “future-land is a region of fairytale” (1977: 53). A Geachian's non-eternalist options are *presentism*, which holds that only *present* objects exist, and the *growing-block theory*, which holds that both past and present (but no future) objects exist.

(37) At any rate, those who maintain that alternative possibilities (i.e. that we “can do otherwise”) are required for freedom (and that we are indeed free) cannot balk at the notion that something that *will* happen *can* fail to happen, unless he or she is an open futurist, i.e. unless he or she denies that there *are* truths about what we will freely do.

(38) Perhaps the Ockhamist would wish to claim that what is objectionable about Geach's view is that what will happen *can epistemically* still fail to happen, though not its claim that what will happen *can causally* still fail to happen. Again, however, while this may be the Ockhamist's official position, I am skeptical that intuition clearly distinguishes between these two senses of “can” at the level of an initial reaction to Geach's position. Is it really *just* the claim that what will happen can *epistemically* fail to happen that is (allegedly) counterintuitive about Geach's proposal? Again, I simply note that philosophers unsympathetic to Geach's view *and* to fatalism ought to tread carefully.

---

(39) Note: there is arguably no contradiction in saying, as would Plantinga and other Ockhamists, both that the future cannot change *and* that what will happen can fail to happen. For Ockhamists, while we cannot change the future, we *can* so act that the actual future *wouldn't have been* the future. In that sense what will happen *can* fail to happen. But the future cannot change: nothing that will happen can become such that it will not.

(40) Though *Providence and Evil* contains the most complete statement of his view, the view is also articulated in Geach (1973). Moreover, in an essay in his 1998 book, *Truth and Hope,* Geach reiterated his view, saying he is “sharply opposed to a view widespread in our day: that…‘changing the future’ is a self-contradictory concept: ‘It was going to happen but didn't happen’ is on this view a violently improper use of language and in no circumstances can be literally true.” See p. 88 of ch. 6, “Prophecy.”

(41) Pike, reprinted in Fischer (1989: 63).

(42) See Adams and Plantinga in Fischer (1989).

(43) Note: the Ockhamist does not claim that *all* soft facts fail to be fixed; just those involving our free decisions.

(44) Geach nowhere says this explicitly, which is somewhat disconcerting, since this is the plain consequence of the account of the mutability of the future he offers. And Geach emphasizes that God must believe all and only what is true.

(45) I owe this point to Neal Tognazzini.

(46) The reason for the past-tensed construction is the following. The Geachian need not—and should not—maintain that “Jones will do X at $t10$” does not entail “Jones (tenselessly) does X at $t10$.” Rather, the Geachian should maintain that such an entailment holds, but that the tenseless “does” must itself be analyzed in terms of tensed propositions, i.e. that “Jones does X at $t10$” is true if “Jones has done X at $t10$, is doing X at $t10$, or will do X at $t10$” is true. Clearly, then, if “Jones will do X at $t10$” is true, so is “Jones does X at $t10$,” since the third disjunct of the analysis will be true. More generally, the Geachian plainly ought to deny that there is some immutable, tenseless fact about what “happens” at times.

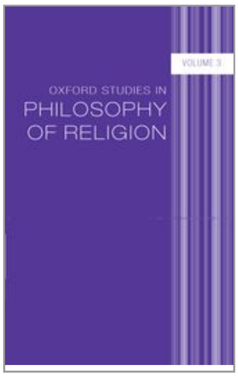
(47) Of course, this is hardly a cost of the view, and many would find such a result to speak in its favor.

(48) Of course, if the reader will permit me a sociological remark, I am not confident that the usage of the term “open theism” in popular discourse will be sensitive to the variety of views that open theists actually (or could possibly) hold. In popular discourse, open theism is often simply the view that God lacks “comprehensive” or “exhaustive” foreknowledge. As Rhoda notes, such a characterization may be appropriate for views like Hasker's, but it is not appropriate for those views denying that future contingents are true. On that view, God *does* have comprehensive or exhaustive knowledge of everything that will be the case. Likewise on Geachianism.

(49) See Hasker (1989).

(50) Other open theists—both of the open-future and limited-foreknowledge variety—might complain that their views can also properly accommodate the providential usefulness of knowing what will be. On such views, God will know that something will be as a result of seeing that it is causally determined. Thus, so long as the future is to some extent causally open, God would be able to use his knowledge of these facts to casually affect other (undetermined) aspects of the future, thereby giving it a more determinate shape. But it is unclear whether this account can truly capture the relevant intuitions. For it seems as if (on this account) what is useful to God is not knowledge of what will be *qua* knowledge of what will be, but knowledge of the causal structure of the world and what it makes inevitable. In other words, that what is causally determined also *will be* is *epiphenomenal* in the relevant sense. For the Geachian, knowing what will be *qua* knowing what will be is by itself useful. At any rate, if other open theists *can* properly accommodate the providential usefulness of such knowledge, then the above claims support merely the *disjunction* of Geachianism and other open views.

(51) For helpful comments on earlier versions of this paper, I wish to thank William Hasker, Eleonore Stump, Alan Rhoda, Kevin Timpe, Michael Nelson, Justin McBrayer, Andrew Moon, Kenneth Boyce, and my colleagues and fellow members of the agency reading group at the University of California, Riverside—Justin Coates, Garrett Pendergraft, Chris Franklin, and especially Philip Swenson and Neal Tognazzini, both of whom were invaluable discussion partners while writing this paper. I am indebted to Mark Hinchliff both for bringing the view to my attention and for several helpful conversations about these topics. Most of all, I owe deep thanks to John Martin Fischer for his constant support and encouragement of this project and for valuable comments on numerous drafts of the paper.



Oxford Studies in Philosophy of Religion Volume 3

Jonathan L. Kvanvig

Print publication date: 2011
Print ISBN-13: 9780199603213
Published to Oxford Scholarship Online: May 2011
DOI: 10.1093/acprof:oso/9780199603213.001.0001

(p.253) Name Index

Adams, Robert Merrihew, 42n, 110–1, 133–6
Almeida, Michael J., 12n
Aquinas, Thomas, 68
Aristotle, 148
Aspect, Alain, 146n
Balog, Katalin, 112n, 115, 129
Barrow, John D., 138n
Bayes, Thomas, 17, 26–31, 36, 139, 188, 202, 208–9
Bell, John Stewart, 146n
Belnap, Nuel, 225
Benacceraf, Paul, 201, 215
Bigelow, John, 73n, 76n
Block, Ned, 114n, 115, 117n, 129
Bohm, David, 160
Boucher, David, 104n
Boyce, Kenneth, 231, 250n
Brandon, Robert, 197n
Brouwer, Luitzen Egbertus Jan, 8
Buller, David, 73n
Cameron, Ross, 1n, 11n, 13n
Carlson, Erik, 42n
Carnap, Rudolf, 216
Carson, Scott, 197n
Carter, Brandon, 138n
Chalmers, David, 114n, 134n
Clarke, Samuel, 167
Clifford, W.K., 190
Clinton, Hillary, 90
Coates, Justin, 250n
Coffman, E.J., 48n, 52n
Collins, Gail, 90

Collins, Robin, 138–9n
Colyvan, Mark, 139n, 187n, 207n, 209–11, 215n
Comesana, Juan, 187
Copp, David, 107
Courant, Richard, 205n
Cowton, Christopher, 90–3, 108
Craig, William Lane, 150–2
Crisp, Roger, 90–3, 108
Crow, James F., 187n, 198n
Darwin, Charles, 56–8, 86, 198–9, 218
Davidson, Donald, 27
Dawid, David, 26, 28
Dembski, William A., 199
Dennett, Daniel, 189
Divers, John, 2n
Dorato, Mauro, 155–6
Dowe, Phil, 59n, 60n, 63–4n, 230n
Dretske, Fred, 112n, 188n, 203n
Driver, Julia, 107n
Duhem, Pierre, 203, 216
Dummett, Michael, 229n
Earman, John, 28
Easwaran, Kenny, 205n
Edwards, A.W.F., 202n
Edwards, Jonathan, 170
Eleonore, Stump, 60n, 250n
Everett, Theodore J., 217n
Field, Hartry, 187n, 201
Finch, Alicia, 42n
Fischer, John Martin, 222n, 229n, 237n, 241n, 250n
Flanagan, Caitlin, 90
Flint, Thomas, 61n
Foot, Philippa, 107–8
Franklin, Chris, 250n
Freddoso, Alfred J., 63n, 229n, 236n
Friedman, Michael, 206–7, 214
Funkhouser, Eric, 13n
(p.254) Garcia, Laura, 7n
Garfield, Jay, 139n
Gauthier, David, 104n
Geach, Peter, 222–250
Gertler, Brie, 114n,
Gillies, Donald, 26, 28
Ginet, Carl, 46
Gödel, Kurt, 145, 162–3
Godfrey-Smith, Peter, 73n, 76
Gonzalez, Guillermo, 163
Goodale, Mel, 130n
Gore, Al, 90

Gott, J. Richard, 150–1
Greco, John, 1n
Green, Mitchell, 225
Greene, Brian, 146
Griffiths, Paul, 73n
Guleserian, Theodore, 4
Gunn, James, 150–1
Hacking, Ian, 202n
Hall, Ned, 163–4
Hampshire, Stuart, 46
Hardy, Thomas, 96
Hare, R.M., 95n
Hart, H.L.A., 46, 65n
Hartshorne, Charles, 224
Hasker, William, 71n, 224n, 225, 245n, 248–9, 250n
Hawthorne, John, 12n
Heller, Mark, 2
Hellman, Geoffrey, 187n, 208–11, 215
Hempel, Carl, 187, 202, 216
Heyer, George S., 16n
Hinchliff, Mark, 237, 250n
Holder, Rodney, 28, 32–3
Honoré, Tony, 65n
Horgan, Terry, 111n, 118n, 119–20
Howard, Mary, 109n
Howard-Snyder, Daniel, 109n
Howard-Snyder, Frances, 102n
Hudson, Hud, 12n, 103n, 109n
Hull, David, 218
Hume, David, 17–27, 32, 34, 36–7, 39, 60n, 64, 174–5
Hunt, David, 41, 44, 47–8
Hurka, Thomas, 107n, 108
Hursthouse, Rosalind, 107–8
Ide, H.A., 58n
Ishiguro, Makio, 205n
Jablonka, Eva, 195n
Jackson, Frank, 114n, 159–60
James, William, 190, 200
Johnston, Mark, 121n
Jung, Carl, 58
Kant, Immanuel, 16n, 98, 103–4
Kapitan, Tomis, 41, 44–8, 50n
Kelly, Paul Joseph 104n
Kennedy, Ted, 90
Kenny, Anthony, 229n
Kimble, Kevin, 111n
Kimura, Motoo, 198n
Kitagawa, Genshiro, 205n
Koons, Robert C., 16n

Kostelecky, V. Alan, 147n
Kotzen, Matthew, 188n, 208
Kripke, Saul, 114n, 179n
Kvanvig, Jonathan L., 1n, 11n, 62n, 223, 229–236, 249
La Croix, Richard R., 41n, 49n
Ladyman, James, 153, 154n, 161n
Lamb, Marion J., 195n
Leibniz, Gottfried, 58–62, 136–6, 166
Lenski, Richard E., 193
Leslie, John, 171
Levin, Michael, 80–1n
Lewens, Tim, 73n
Lewis, David, 3, 7n, 12n, 60n, 61, 161n, 162, 231
Lewontin, Richard, 187
Loar, Brian, 112n, 114n, 115
Lockwood, Michael, 125n
(p.255) Lorentz, Hendrik, 147
Lucas, John Randolph, 224
Lyon, Aidan, 215n
Maddy, Penelope, 205n, 208n, 213–4, 217n
Malebranche, Nicolas, 62
Manson, Niel, 139n
Markov, Andrey, 190
Maudlin, Tim, 147n
Mayr, Ernst, 196
Mendel, Gregor, 197
Monod, Jacques, 187, 216
Monton, Bradley, 146n
McBrayer, Justin, 250n
McCain, John, 231
McCann, Hugh, 62n
McCarthy, Joe, 91
McGrew, Lydia, 17n, 22–3, 27n
McGrew, Timothy J., 17n, 22–3, 27n
McLaughlin, Brian, 112n, 115–33
McLaughlin, Peter, 73n
Mele, Alfred, 45
Menzel, Christopher P., 1n
Menzies, Peter, 65n
Merricks, Trenton, 49n
Mill, John Stuart, 64
Millican, Peter, 26, 28n
Millikan, Ruth, 73n, 74, 75–7n
Milner, David, 130n
Minkowski, Hermann, 148
Mittler, John E., 193
Molina, Luis de, 60–1, 223
Moon, Andrew, 250n
Morris, Thomas, 4–5, 8

Nahmias, Eddy, 41n
Nanay, Bence, 187
Neander, Karen, 73n, 76n
Nelkin, Dana, 46, 50n
Nelson, Michael, 250n
Newton, Isaac, 148–9, 206–7
Nichols, Peter, 187
Nolan, Daniel, 1n
Nozick, Robert, 20
Numbers, Ronald L., 212n
Obama, Barack, 231
Ockham, William of, 199, 223, 247–9
O'Connor, Timothy, 113n, 135n, 137n, 139n, 166–86
Owens, David, 26–7, 57n
Paley, William, 31
Papineau, David, 112n, 115, 129, 201n
Parfit, Derek, 105n
Pargetter, Robert, 73n, 76n
Parsons, Charles, 207n
Paseau, Alexander, 187
Pendergraft, Garrett, 250n
Pennock, Robert T., 212n
Pereboom, Derk, 50n
Pettit, Philip, 50n
Pigliucci, Massimo, 195n
Pike, Nelson, 223, 240–5
Plantinga, Alvin, 8, 42n, 67–71, 76n, 77–81, 84–6, 190, 198n, 236–241
Plateau, Joseph, 205–6
Polkinghorne, 63n
Popkin, Richard H., 20n
Popper, Karl, 210
Pospelov, Maxim, 147n
Priest, Graham, 139n
Prior, Arthur N., 224n
Provine, Will, 189
Pruss, Alexander, 139n, 175
Putnam, Hilary, 144–5, 147, 151, 162, 189, 207, 209–10, 217
Quine, Willard van Orman, 162, 189, 203, 207, 209–11, 216
Quinn, Philip L., 1, 41n
Randall, John Herman, 18
Reichenbach, Hans, 57, 59, 63, 217n
(p.256) Resnik, Michael, 209–11
Rhoda, Alan, 222–5, 236n, 242, 245, 250n
Richards, Jay, 163
Robbins, Herbert, 205n
Romalis, Michael, 147n
Romney, Mitt, 90
Rosenbaum, Stewart, 1n, 11n
Rosenberg, Alexander, 187

Ross, W.D., 153, 154n, 161n
Rowe, William, 6n, 42n, 150, 174
Royall, Richard, 202n
Ruse, Michael, 187, 189
Russell, Bertrand, 39
Sakamoto, Yosiyuki, 205n
Sarkar, Sahotra, 195n
Sartorio, Carolina, 187
Schindler, Oskar, 94, 96
Schlesinger, George, 26, 32
Schramm, David, 150–1
Schwartz, Peter, 77n
Scotus, John Duns, 167
Senor, Tom, 13n, 42n
Shakespeare, William, 89, 98
Shapiro, Larry, 187
Sheehy, Paul, 11n
Sidelle, Alan, 187
Sider, Theodore, 145, 147–9, 151, 162
Slote, Michael, 107–8
Smart, J.J.C., 114n, 236–8, 249
Smolin, Lee, 146
Sobel, David, 107
Sobel, Howard, 18, 26, 28
Sober, Elliott, 73n, 189n, 190n, 192n, 197n, 205n, 207n, 209, 213n, 216–7n
Soifer, Eldon, 92–3n, 95n
Sorensen, Roy, 18, 32–3
Spinoza, Baruch, 136, 212
Stalnaker, Robert, 114n, 231
Stanwix, Paul, 147
Steinhardt, Paul, 154
Stephen, Leslie, 18
Suppes, Patrick, 60n
Swenson, Philip, 250n
Swinburne, Richard, 17n, 26, 110–1, 133–6, 199n, 224
Szabados, Béla, 92–3n, 95n
Tienson, John, 111n, 118–20
Timpe, Kevin, 250n
Tinsley, Beatrice, 150–1
Tippler, Frank L., 138n
Titelbaum, Michael, 187
Tognazzini, Neal, 222n, 243n, 250n
Tooley, Michael, 81n, 147n
Turner, Dan, 92
Turok, Neil, 154
Tye, Michael, 112n, 115–7, 120–9, 131–33
Tymoczko, Thomas, 205n
Van Fraassen, Bas, 150, 153
Van Inwagen, Peter, 3n, 41n, 50n, 52n, 55–6, 58, 61n, 160–1, 224n, 225

Van Roojen, Mark, 102n
Wainwright, William J., 42n
Waller, Bruce N., 50n
Warfield, Ted A., 42n, 48n, 49n, 52n
White, Stephen, 114n,
Wright, Larry, 73n